

Running Head: PERCEPTIONS OF TEACHER EVALUATION MODELS

A COMPARISON OF INDIANA ADMINISTRATOR AND TEACHER PERCEPTIONS OF
TEACHER EVALUATION MODELS

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF EDUCATION

BY
BRET DAGHE
DR. KENDRA LOWERY – CHAIR

BALL STATE UNIVERSITY
MUNCIE, INDIANA

JULY 2018

PERCEPTIONS OF TEACHER EVALUATION MODELS

A COMPARISON OF INDIANA ADMINISTRATOR AND TEACHER PERCEPTIONS OF
TEACHER EVALUATION MODELS

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF EDUCATION

BY

BRET DAGHE

DR. KENDRA LOWERY – CHAIR

APPROVED BY:

Committee Chairperson_____
Date_____
Committee Chairperson_____
Date_____
Committee Chairperson_____
Date_____
Committee Chairperson_____
Date_____
Dean of Graduate School_____
Date

BALL STATE UNIVERSITY

MUNICE, INDIANA

JULY 2018

PERCEPTIONS OF TEACHER EVALUATION MODELS

DEDICATION

This dissertation is dedicated to my wife, Laura, and my two children, Jake and JoJo. It is only because they believed in me that kept me going throughout the process. Their love, devotion, and dedication to our family while my time and attention was often fractured with the demands of coursework, research, synthesis, and writing was very valuable to me and will never be forgotten.

PERCEPTIONS OF TEACHER EVALUATION MODELS

ACKNOWLEDGMENTS

Thank you to my Ball State University committee for all their guidance and wisdom through this exciting learning journey. My committee chair, Dr. Kendra Lowery, provided endless hours of support through discussions and feedback every step of the way. My committee members, Dr. Marilyn Quick, Dr. Lori Boyland, and Dr. David Pearson are also to be commended for their support throughout this educational endeavor. This final step in my educational journey is truly something I was never sure I would even attempt, let alone complete.

I also have to take the time to say thanks to all my supporters and colleagues for their ability to listen and process with me when I was at my highest and lowest stages of this process. I thank them for their words of encouragement, any and all additional work that had to be done in my absence, and their lifetime of friendships. I say thank you to Dr. Jim Snapp and Dr. Kat Jessup who pushed me down this educational road only to know that in the long run it would truly pay off. I would like to say thanks to all the educators and administrators at the three schools where I was able to spend time observing and interviewing to get the information to conduct my research.

PERCEPTIONS OF TEACHER EVALUATION MODELS

ABSTRACT

DISSERTATION: A Comparison of Indiana Administrator and Teacher Perceptions of Teacher Evaluation Models

STUDENT: Bret Daghe

DEGREE: Doctor of Education

COLLEGE: Teacher's College

DATE: July 2018

PAGES: 193

This qualitative multiple case study was based on in-depth interviews with participants from three different schools. Each school used a different High-Stakes Accountability Model for teacher evaluation: RISE Evaluation and Development (RISE), Teacher Advancement Program (TAP), and Peer Assistance and Review (PAR). The purpose of this study was to analyze which of the models' characteristics teachers and administrator perceived to be most effective in promoting professional development and student achievement. The Joint Committee on Standards for Educational Evaluation's Utility Standards were used as a conceptual framework. Major findings include: structured evaluation models facilitate professional growth when they include training and feedback and promote positive changes to instruction, and they promote increased student achievement when the process results in more data-informed and student-focused teaching.

Keywords: teacher evaluation models, professional growth, student achievement, RISE Evaluation and Development System, Teacher Advancement Program, Peer Assistance and Review

PERCEPTIONS OF TEACHER EVALUATION MODELS

TABLE OF CONTENTS

A COMPARISON OF INDIANA ADMINISTRATOR AND TEACHER PERCEPTIONS OF TEACHER EVALUATION MODELS.....	I
APPROVAL PAGE	II
DEDICATION.....	III
ACKNOWLEDGMENTS	IV
ABSTRACT.....	V
TABLE OF CONTENTS	VI
LIST OF TABLES	X
LIST OF FIGURES	XI
CHAPTER ONE: INTRODUCTION.....	1
STATEMENT OF THE PROBLEM.....	5
PURPOSE OF STUDY	5
CONCEPTUAL FRAMEWORK.....	6
RESEARCH QUESTIONS.....	9
SIGNIFICANCE OF THE STUDY	9
DEFINITION OF TERMS.....	11
DELIMITATIONS.....	12
ORGANIZATION OF THE STUDY.....	12
CHAPTER TWO: LITERATURE REVIEW.....	14
FEDERAL AND STATE ACCOUNTABILITY ERA IN EDUCATION	14
HISTORY OF TEACHER EVALUATION POLICIES AND PRACTICES.....	16
RESEARCH ON TEACHER EVALUATIONS IN INDIANA	23
CHANGES TO CURRENT PRACTICES AND STANDARDS-BASED EVALUATIONS	25
CURRENT BEST PRACTICES REGARDING TEACHER EVALUATIONS.....	27
<i>Overall framework.....</i>	27
<i>Expectation of quality.....</i>	27
<i>Problems to be overcome</i>	28
<i>Competing views.....</i>	29
<i>Teacher evaluations as a growth model.....</i>	30
<i>Other key components of teacher evaluations.....</i>	32
OVERVIEW OF THE THREE EVALUATION MODELS (RISE, TAP, AND PAR)	34
<i>RISE.....</i>	35
<i>TAP</i>	36
<i>PAR.....</i>	36
SIMILARITIES AND DIFFERENCES AMONG THE THREE MODELS	37
FRAMEWORKS: DANIELSON AND MARZANO	41

PERCEPTIONS OF TEACHER EVALUATION MODELS

SUMMARY	43
CHAPTER THREE: METHODOLOGY	45
RATIONALE AND ASSUMPTIONS FOR THE QUALITATIVE DESIGN.....	45
RESEARCHER’S ROLE	47
SITE AND PARTICIPANT SELECTIONS	47
DATA COLLECTION TECHNIQUES	50
MANAGING AND RECORDING DATA	52
DATA ANALYSIS PROCEDURES.....	53
LIMITATIONS	55
SUMMARY	56
CHAPTER FOUR: FINDINGS – CASE STUDY #1.....	57
WASHINGTON HIGH SCHOOL: PROFILE	57
WASHINGTON HIGH SCHOOL: RISE EVALUATION MODEL	59
WASHINGTON HIGH SCHOOL: PARTICIPANTS.....	61
WASHINGTON HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 1	62
<i>Theme 1: RISE caused changes to instruction</i>	<i>62</i>
<i>Theme 2: RISE generated valuable feedback.....</i>	<i>64</i>
<i>Theme 3: RISE included beneficial training.....</i>	<i>66</i>
<i>Theme 4: Timing of certain RISE components affected timing of growth.....</i>	<i>68</i>
WASHINGTON HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 2	68
<i>Theme 1: Teaching became more student-focused.....</i>	<i>69</i>
<i>Theme 2: Data used to benefit students.....</i>	<i>70</i>
WASHINGTON HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 3	71
<i>Theme 1: Most teachers looked favorably on RISE</i>	<i>71</i>
<i>Theme 2: Negative perceptions did not influence overall processes or outcomes</i>	<i>72</i>
WASHINGTON HIGH SCHOOL: SUMMARY	75
CHAPTER FIVE: FINDINGS – CASE STUDY #2	77
JEFFERSON HIGH SCHOOL: PROFILE	77
JEFFERSON HIGH SCHOOL: TAP EVALUATION MODEL	79
JEFFERSON HIGH SCHOOL: PARTICIPANTS	83
JEFFERSON HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 1	84
<i>Theme 1: Training under TAP led to growth</i>	<i>84</i>
<i>Theme 2: TAP caused positive changes to instruction.....</i>	<i>86</i>
<i>Theme 3: Feedback from post-conferences and master and mentor teachers caused growth</i>	<i>87</i>
JEFFERSON HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 2	90
<i>Theme 1: Lack of student data.....</i>	<i>90</i>
<i>Theme 2: TAP created a more student-centered classroom.....</i>	<i>91</i>
JEFFERSON HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 3	93
<i>Theme 1: Teachers have gotten more comfortable with TAP</i>	<i>93</i>
<i>Theme 2: Buy-in needed to use TAP to improve</i>	<i>94</i>

PERCEPTIONS OF TEACHER EVALUATION MODELS

JEFFERSON HIGH SCHOOL: SUMMARY	95
CHAPTER SIX: FINDINGS – CASE STUDY #3	98
LINCOLN HIGH SCHOOL: PROFILE	98
LINCOLN HIGH SCHOOL: PAR EVALUATION MODEL	99
LINCOLN HIGH SCHOOL: PARTICIPANTS	102
LINCOLN HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 1	102
<i>Theme 1: Positive changes to instruction.....</i>	<i>103</i>
<i>Theme 2: Feedback under PAR promoted growth</i>	<i>104</i>
<i>Theme 3: Training facilitated growth.....</i>	<i>108</i>
<i>Theme 4: PAR provides common and clear language for shared expectations.....</i>	<i>109</i>
LINCOLN HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 2	111
<i>Theme 1: Improved instructional strategies.....</i>	<i>111</i>
<i>Theme 2: PAR required use of student achievement data increased student achievement. .</i>	<i>113</i>
LINCOLN HIGH SCHOOL: THEMES THAT EMERGE FROM THE RESEARCH QUESTION 3	116
<i>Theme 1: PAR was liked, so followed.....</i>	<i>116</i>
<i>Theme 2: Concern about timing</i>	<i>118</i>
LINCOLN HIGH SCHOOL: SUMMARY	119
CHAPTER SEVEN: ANALYSIS OF THEMES	121
THEMES PERTAINING TO PROFESSIONAL GROWTH.....	123
<i>Theme #1: Evaluation models facilitate professional growth when they include training</i>	<i>123</i>
<i>Theme #2: Evaluation models facilitate professional growth when they promote positive</i>	<i>125</i>
<i>change to instruction</i>	<i>125</i>
<i>Theme #3: Evaluation models facilitate professional growth when they include feedback</i>	<i>127</i>
THEMES PERTAINING TO STUDENT ACHIEVEMENT	129
<i>Theme #1: Evaluation models promote an increase in student achievement when they</i>	<i>129</i>
<i>support teaching to be more student focused</i>	<i>129</i>
<i>Theme #2: Evaluation models promote an increase in student achievement when they</i>	<i>132</i>
<i>incorporate student data.....</i>	<i>132</i>
THEME PERTAINING TO THE INFLUENCE OF PERCEPTIONS ON PROCESS AND OUTCOMES	134
<i>Evaluation models that are perceived positively are more likely to be used by teachers, and</i>	<i>134</i>
<i>thus, have a greater influence on processes and outcomes</i>	<i>134</i>
SUMMARY	137
CHAPTER EIGHT: CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS	138
SUMMARY OF THE STUDY	138
<i>Overview of the problem.....</i>	<i>138</i>
<i>Research questions</i>	<i>139</i>
<i>Review of the methodology.....</i>	<i>140</i>
<i>Major findings</i>	<i>141</i>
<i>Findings related to the literature.....</i>	<i>141</i>

PERCEPTIONS OF TEACHER EVALUATION MODELS

<i>Research Question 1: What characteristics of teacher evaluation models do teachers and administrators perceive to be the most effective in promoting professional growth?.....</i>	<i>143</i>
<i>Research Question 2: What characteristics of teacher evaluation models do teachers and administrators perceive to be the most effective in promoting student achievement?</i>	<i>147</i>
<i>Research Question 3: In what ways do teacher and administrator perceptions of the evaluation models influence the evaluation process and educational outcomes?</i>	<i>149</i>
UNEXPECTED FINDINGS	152
CONCLUSIONS	153
<i>Implications for practice</i>	<i>154</i>
RECOMMENDATIONS FOR FURTHER RESEARCH.....	156
CONCLUDING REMARKS	157
REFERENCES.....	158
APPENDIX A: INTERVIEW QUESTIONS FOR ALL PARTICIPANTS.....	174
APPENDIX B: RECRUITMENT SCRIPT FOR POTENTIAL TEACHER PARTICIPANTS.....	176
APPENDIX C: RECRUITMENT SCRIPT FOR POTENTIAL ADMINISTRATOR PARTICIPANTS.....	177
APPENDIX D: INFORMED CONSENT FOR TEACHER PARTICIPANTS.....	178
APPENDIX E: INFORMED CONSENT FOR ADMINISTRATOR PARTICIPANTS...	180

PERCEPTIONS OF TEACHER EVALUATION MODELS

LIST OF TABLES

TABLE 1 SIMILARITIES AND DIFFERENCES WITH THE THREE MODELS	39
TABLE 2 COMPARISON OF THE THREE HIGH SCHOOLS	48
TABLE 3 COMPARISON OF ADMINISTRATORS AND TEACHERS	50
TABLE 4 DEMOGRAPHIC BREAKDOWN OF WHS	58
TABLE 5 DEMOGRAPHIC BREAKDOWN OF JHS	78
TABLE 6 DEMOGRAPHIC BREAKDOWN OF LHS	99
TABLE 7 COMPARISON OF KEY DATA IN WHS, JHS, AND LHS MODELS	122

PERCEPTIONS OF TEACHER EVALUATION MODELS

LIST OF FIGURES

FIGURE 1 WEIGHTING OF TAP RUBRIC DOMAINS AT JHS	82
FIGURE 2 PERCENTAGES USED TO CALCULATE CUMULATIVE OR SUMMATIVE PERFORMANCE SCORES AT JHS	83

PERCEPTIONS OF TEACHER EVALUATION MODELS

CHAPTER ONE: INTRODUCTION

In this time of heightened educational accountability for school performance, the outcomes for teacher evaluations have never been more scrutinized nor have unprecedented expectation for the results of these evaluations been made more public (Knoeppel & Reinhart, 2008; Ravitch, 2010; Timperly, 2011; Zhao, 2009). Beginning with the report *A Nation at Risk* (National Commission on Excellence in Education, 1983), the topic of effective teacher evaluations became a greater part of the education discussion with politicians. The social and educational impacts caused by the report led to concerns that the American education system was no longer internationally competitive (Burke & Marshall, 2010). This caused a pressing need for national uniformity of federally imposed accountability requirements to improve teacher performance (Gutherie & Springer, 2004; Taylor & Tyler, 2012; Zeichner, 2011). In 2001, Congress reauthorized the Elementary and Secondary Education Act (ESSA) through the No Child Left Behind Act (2002). In 2002, under the George W. Bush administration, the federal government passed accountability standards for teachers calling for the focus to be on teacher subject specific knowledge for teacher credentialing (Hursh, 2007; Paige, 2002).

Under the No Child Left Behind (NCLB) law, all schools were required to improve the educational environment and implement new standards and assessments intended to categorize educational excellence. This federal mandate significantly increased the federal role in holding schools responsible for the academic progress and improvement of student learning of all students (Hursh, 2007; Lazaridou & Iordanides, 2011; Superfine, Gottlieb & Smylie, 2012). The law also required states to ensure their teachers were “highly qualified,” which meant teachers had to have a bachelor’s degree in the subject they were teaching and hold state certification

PERCEPTIONS OF TEACHER EVALUATION MODELS

(Paige, 2002; U.S. Department of Education, 2004). This increased involvement caused many to say that:

[u]nder NCLB, the federal government's role has become excessively intrusive in the day-to-day operation of public education by trying to incorporate the principles of individual state standards-based reforms and condensing them in one federal statute that imposes a one-size-fits-all accountability system.

(National Conference of State Legislatures, 2005, p.11)

NCLB substantially increased the testing requirements for states and established demanding accountability standards for schools, districts, and states with measurable adequate yearly progress (AYP) objectives for all students. Consequently, the percentage of students who scored at the proficient level or higher on the state assessments varied widely from state to state, which in turn led to a focus on meeting unrealistic goals of 100% student proficiency, instead of measuring progress and growth (Lee & Orfield, 2006). The effects of NCLB and the one-size-fits-all measuring student achievement and providing consequences for missing goals were significant to schools and states (Barnes & Slate, 2013; Crawford, 2004; Sawchuk, 2010; Selwyn, 2007).

In 2009, the federal Race to the Top (RTTT) initiative placed educator accountability on the national level and further caused changes to the teacher evaluation structure (Civic Impulse, 2016). To an even greater magnitude, RTTT raised the bar on public awareness and sentiment about principal and teacher evaluation practices to the extent that states and, therefore schools, were required to implement teacher evaluation that would focus on effective instruction and annual student growth for all staff members (McGuinn, 2012; Olivia, Mather, & Laine, 2009; U.S. Department of Education, 2009). In addition, the financial lure of potential federal dollars

PERCEPTIONS OF TEACHER EVALUATION MODELS

raised the ante for states to implement new teacher evaluation systems (Hazi & Rucinski, 2009). Given the potential for increased federal dollars, nearly half of the states voluntarily submitted proposals and were awarded the RTTT funds (Onosko, 2011). Even though there were states that did not apply for, nor were granted, the RTTT dollars, the expectations of desired changes in education practice, especially teacher evaluation, did not diminish. As a result, the RTTT program had a significant impact on the national political discourse around education and pushed many states, including Indiana, to enact policy changes and laws around the need to improve and change the teacher evaluation process (Mathis, 2011; McGuinn, 2012).

As a direct result of RTTT, Indiana lawmakers passed Public Law 90 (P.L. 90) legislation in 2011, which centered on educator effectiveness and required more stringent reviews of teacher performance (IC 20-28-11.5, 2011). The law required every school corporation to develop a plan to complete yearly teacher evaluations and the implementation of the plan to begin in the fall of the 2012-2013 school year. This law mandated that all licensed teachers, administrators, and staff members be evaluated annually in every Indiana school district. However, the law did, in fact, provide some flexibility to school districts to select the type of evaluation model to utilize, as long as the model chosen incorporated analysis of student growth data (Baker, Oluwole, & Green, 2013). Every evaluation model used by Indiana school districts must incorporate the four categories: Highly Effective, Effective, Improvement Necessary, and Ineffective (Baker et al., 2013). The final requirement and perhaps the most controversial component of the law was that evaluations must be linked to the pay raises and job security of Indiana teachers and administrators.

Although many requirements and expectations were mandated, the directive as to what evaluation model school districts were to use was not specified. Every Indiana school district

PERCEPTIONS OF TEACHER EVALUATION MODELS

was required to evaluate all personnel by whatever model school leaders thought would meet district needs. This created a system of teacher evaluations where various school districts could use a wide variety of tools creating the potential of districts measuring different aspects of instruction. Three different evaluation models came to the forefront in Indiana: the RISE evaluation and development system (RISE), the Teacher Advancement Program (TAP), and the Peer Assistance and Review (PAR) (IC 20-28-11.5, 2011).

RISE is the model plan developed by the Indiana Department of Education (IDOE) and piloted in 2011-2012. It was touted as being able to “paint a fair, accurate, and comprehensive picture of a teacher’s performance” by evaluating the two major components: professional practice (teacher instructional knowledge and skills) and student learning (IDOE, 2012, p.3). TAP and PAR are pre-developed models (1999 and early 1980s respectively) that are used nationally. TAP uses a clearly defined, five-point scale rubric for evaluation of teachers by classroom observation and measures student performance based on achievement gains. PAR involves teachers and administrators working together to develop and implement a teacher evaluation protocol.

The next section presents information to analyze the problem addressed by the research and supplies a brief summary of the most relevant research and theory pertaining to the subject of this qualitative study. This qualitative case study permitted the exploration of teacher evaluation models from the perceptions of teachers and administrators. Yin (2013) explained that a case study enables researchers to conduct an exploration from an angle that is both holistic and real-world.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Statement of the Problem

The problem is that schools are complex, dynamic systems that influence students' academic success, and the evaluation models they employ need to be constantly scrutinized from multiple perspectives to determine if teachers and administrators perceive the evaluations to enhance their ability to promote professional growth and improve instruction. According to Whiteman, Shi, & Plucker (2011), prior to the implementation of law P.L. 90, which mandated that teacher evaluations be done yearly on every school employee, not every educator in the State of Indiana was evaluated on an annual basis. In fact, often times years went by between evaluations, especially evaluations conducted for tenured teachers. Teacher evaluation reform has clearly happened over the last five years. What remained to be answered was whether this current course of action had changed the teacher evaluation process from the teacher and administrator perspective. Here, I studied three different teacher evaluation models (RISE, TAP, and PAR) to understand the differences in the effectiveness of each of the models for promoting professional growth and teacher effectiveness. This was important because research indicates that well-designed teacher evaluation models have a direct and lasting effect on individual teacher performance (Danielson & McGreal, 2000; Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Ellett & Teddlie, 2003; Guskey, 2000; Marzano & Toth, 2013; Taylor & Tyler, 2012).

Purpose of Study

The purpose of this study was to analyze teacher and administrator perceptions regarding three High-Stakes Accountability Models— RISE Evaluation and Development System (RISE), Teacher Advancement Program (TAP), and Peer Assistance and Review (PAR). Data was collected on teacher and administrator perceptions concerning: (a) the general framework and

PERCEPTIONS OF TEACHER EVALUATION MODELS

structure of the three teacher evaluation models; (b) the specific practices for gathering and compiling the evaluation data used in each of the three models (i.e., the forms used); and (c) the potential impact of each model on teacher effectiveness, student achievement, and professional growth. Using this collected data, I compared and contrasted the three different models in a qualitative case study.

Conceptual Framework

The six major assumptions that guided the development of the Personnel Evaluation Standards developed by The Joint Committee on Standards for Educational Evaluation (2009) were used as the foundation for the conceptual framework of this study. These six assumptions were classified as “must haves” in all evaluation models. In other words, the assumptions were the driving force in determining the components that every teacher evaluation needed to have in order to be successful. They also represent a belief by educational practitioners that when the Personnel Evaluation Standards are met, the evaluation will enhance the fairness and quality of teacher growth and professional practice (Joint Committee on Standards for Educational Evaluation, 2009). The six assumptions are as follows:

- 1) The fundamental purpose of personnel evaluations must be to help provide effective services to students;
- 2) Personnel evaluation practices should be constructive and free of unnecessarily threatening or demoralizing characteristics;
- 3) Personnel evaluations should be conducted with an understanding that the evaluation system adheres to culturally competent practices;
- 4) Personnel evaluations are vital for planning sound professional development and training experiences;

PERCEPTIONS OF TEACHER EVALUATION MODELS

- 5) Personnel evaluation standards allow schools to define their own expectations, approaches, and role definitions to meet performance and accountability standards; and
- 6) Personnel evaluations must provide more time and effort devoted to applying standards when personnel actions are uncertain or potentially controversial.

These six assumptions aided in the organization of the 27 Personnel Evaluation Standards into categories defined by four essential attributes: propriety, utility, feasibility, and accuracy (JCSEE, 2009). Of the four attributes, the utility category was most directly applicable to my research regarding the effectiveness of the RISE, TAP, and PAR models for promoting teacher growth, and the six standards within that category (the Utility Standards) are what I focused on.

An explanation of the six Utility Standards is helpful to gain a better understanding of their relation to this study. In general, the Utility Standards “view personnel evaluation as an integral part of an institution’s ongoing effort to . . . provide [staff members] with timely and relevant evaluative feedback and encourage and guide them to deliver high quality service” (JCSEE, 2009, p. 6). The first Utility Standard probes the question of whether the personnel evaluation is constructive. This means it reflects the institution’s goals and missions and builds evaluatees’ “professional self-knowledge, increase[s] their enthusiasm, and their efficacy as practitioners” (p. 70). The second Standard requires evaluators to consult, at the start of the evaluation process, with the intended user groups to determine the intended uses of the evaluation results and findings so that evaluations will be guided by their intended uses. The third Standard deals with personnel evaluations having to be completed by credible, well-trained evaluators. The expectation is that the evaluator has the title to evaluate, be trained or certified, and be able to understand the expectations of the position being evaluated. The fourth Utility Standard states that personnel evaluations need clear performance expectations in words that are

PERCEPTIONS OF TEACHER EVALUATION MODELS

easily understood by the evaluator and evaluatee. This means that the evaluation process has job expectations and a rating scale or rubric that are clearly defined and explicitly aligned. The fifth Standard requires personnel evaluations to provide clear and timely results so professional development and personal growth can happen effectively. Thus, all evaluations should result in clearly written reports that accurately measure evaluatees' job performance. The last Utility Standard involves personnel evaluations being used to improve staff performance as a whole. This means that evaluation process leads to individualized and full staff professional development plans and programs.

In comparing and summarizing the RISE, TAP, and PAR evaluation models, at least on a surface level, all three models seem to address these proper teacher evaluation standards. My study focused on whether what the models claim or intend to address equate to how the models are actually used in practice in school districts across Indiana. Of the possibilities I might have encountered, I wondered if I would find that administrators and teachers report that in order for teacher growth to occur, the evaluation system must provide teachers with detailed feedback, heightened awareness of improved pedagogies, and assistance in the implementation of useful teaching techniques (Goe, Biggers, & Croft, 2012; Humphrey, Koppich, Bland, & Bosetti, 2011).

Research indicates that teacher evaluations need to be fair and conclusions reached must accurately reflect the teacher's actual level of performance (Alexander, 2016; Danielson, 2012). Teacher evaluations, by their very nature are difficult, too often done quickly with desired outcomes not being reached and are seldom done well with positive results (Weisberg, Sexton, Mulhern, Keeling, Schunck, Palcisco, & Morgan, 2009). It is essential to identify what current practice is so that practitioners may make research-based adjustments to their practice if needed. Therefore, the next section includes my three research questions for this study.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Research Questions

The research questions allowed me to gather information about teacher and administrator perceptions of the three different evaluation models over the last five years. I compared teachers' and administrators' perceptions of the extent to which these evaluation models increase support for teacher effectiveness, raise student achievement accountability, and bring about desired professional growth by Indiana teachers and administrators.

The research questions are:

1. What characteristics of teacher evaluation models do teachers and administrators perceive to be the most effective in promoting professional growth?
2. What characteristics of the teacher evaluation models do teachers and administrators perceive to be the most effective in promoting student achievement?
3. In what ways do teacher and administrator perceptions of the evaluation models influence the evaluation process and educational outcomes?

Significance of the Study

A review of the literature revealed that the majority (seven of nine studies) of the research conducted post-P.L. 90 pertaining to teacher evaluations in Indiana has used a quantitative research methodology. Thus, this qualitative study will add to this underrepresented category. Administrators have the professional responsibility and accountability to focus on teacher growth and effectiveness (Hill, Charalambous, & Kraft, 2012; Halverson & Clifford, 2006). Teaching and learning are at the core of educational practice, and teacher quality is a very important school-level factor affecting student achievement (Goldhaber, Goldschmidt, & Tseng, 2013). This study examines areas of individual teacher evaluations and whether they have been conducted with fidelity, according to the rules of the evaluation model; methods and

PERCEPTIONS OF TEACHER EVALUATION MODELS

tools used in evaluations and whether teachers and administrators perceive them to have been effective; and professional development and its ties to evaluations (Danielson, 2012; Kane & Stager, 2008). Teacher evaluation models are developed using test scores, and they are used for decisions regarding employment and compensation which can be averse to the teacher evaluation process (Baker et al., 2010; Burris & Welner, 2011).

This study was conducted to determine perceptions of how three evaluation models—RISE, TAP, and PAR—have an effect on professional growth for teachers. Using administrator and teacher perceptions, I analyzed the various practices such as scheduled evaluations versus unscheduled evaluations, the number of evaluations, the timing of evaluations, the type of feedback given, whether students and/or parents should participate in the evaluations, and whether consulting teachers should participate. This research will contribute to the existing body of literature focused on effective practices within each of the three teacher evaluation models and will be beneficial given the potential for upcoming changes to Indiana teacher evaluation systems. Future elected officials, committee members, and leaders within the Indiana Department of Education can use the data from this study to determine if the current models are successful in their purpose. They can also use the data from this study on teacher and administrator perceptions to determine what changes need to be made to the models in order to improve teacher effectiveness.

Also, as states across the country have passed legislation mandating teacher evaluation reform, local districts have been called upon to implement new systems for evaluating the effectiveness of their teachers (McGuinn, 2012; Olivia et al., 2009; USDOE, 2009). Traditionally, in many local school districts, evaluation systems were subject to collective bargaining and, therefore, highly influenced by teacher unions. However, several states have

PERCEPTIONS OF TEACHER EVALUATION MODELS

removed teacher evaluation from the scope of collective bargaining to give district administrators the freedom to implement expansive teacher evaluation reforms (Blume, 2010; Shaha, Glassett, & Copas, 2015; Spring, 2010; Superfine, 2016). Thus, while not the focus of my research, I was open to determining how collective bargaining has affected current teacher evaluation practices. So, as I gathered school-level data, evidence that collective bargaining influenced teacher evaluations will necessarily cause me to adapt my questions to clarify the scope of that influence; e.g., seeking the nature of the bargained for language, learning the process of negotiations, identifying the influencers in the process.

Results from the study may provide educators in Indiana with relevant information that can be used to further enhance evaluation techniques and models. I hoped to discover that both the administrators and teachers have seen growth in the teacher evaluation process since the inception of P.L. 90 five years ago and the use of structured models for evaluation.

Definition of Terms

The following definitions are offered to provide an understanding of key terms used in this study:

- *Formative evaluation* – normal and informal evaluation periods throughout the school year of teachers by administrators who observe the teachers' classroom teaching and provide feedback focused on improvement (Popham, 2013)
- *Summative evaluation* – a final evaluation of teachers by administrators that is based on the compilation of data collected during the formative evaluations and is used to monitor overall progress toward expectations with an end in mind to make personnel decisions (Popham, 2013)

PERCEPTIONS OF TEACHER EVALUATION MODELS

- *Teacher effectiveness* – the degree to which teachers meet the expectations of the school district in their classrooms, taking into account the areas assessed in formative and summative evaluations (Darling-Hammond, 2013)
- *Teacher evaluation* – the collection and use of information to measure against an established criteria resulting in a judgment regarding teacher performance made by an administrator (Danielson & McGreal, 2000)
- *Teacher evaluation system* – the comprehensive methodology and structure used by a school district in order to use teacher evaluations to measure, critique, and provide feedback regarding the job performance of the district's teachers (Marzano & Toth, 2013)

Delimitations

This study did not look at the value-added measures of teacher evaluations. Value-added measures involve teachers being evaluated based on student improvement on test scores from year to year. This was not included because none of the three evaluation models studied include value-added measures as a component of teacher evaluations. In addition, the scope was limited to the state of Indiana because evaluation systems are handled by each state individually and this was not a cross-state comparison. Finally, the number of participants was limited to nine individuals (three administrators and six teachers) at three school districts so that I could best make use of the qualitative methodology and explore the individual cases in appropriate and meaningful depth.

Organization of the Study

The study was organized into eight chapters. Chapter 1 gives an overview of the changes to teacher evaluation expectations and the accountability measures that are now in place with new federal and state laws. This chapter lays the groundwork for the significance of the study

PERCEPTIONS OF TEACHER EVALUATION MODELS

and why I am choosing to do this study. It is important to understand the perceptions of teacher and administrators to understand best practices with regard to teacher evaluations. The data collection also allowed for important information to be gathered and utilized for future development of teacher evaluation models. Chapter 2 contains a review of the current literature and studies that have centered on teacher evaluations. This study will be unique because it compares the feedback of administrators and teachers to gain a better understanding between what has happened with evaluations currently and what changes might be considered in future evaluation models. Chapter 3 outlines the research questions and describes the research design. Furthermore, this chapter will give a description of the sample, the instrument used, and how the data will be collected and analyzed. This chapter ends with a listing of the limitations of the study and concluding remarks. Chapters 4, 5, and 6 outline the results of the administrator and teacher interviews at each of the three schools. Chapter 7 analyzes the themes and patterns developed from the interviews, analysis of the evaluation tools under each model, and general observations made during the interviews outlined in Chapters 4-6. Finally, Chapter 8 summarizes the conclusions, implications, and recommendations for this qualitative case study. The next section of the study is the literature review, which provides an overview of the research that has been conducted in the area of teacher evaluations.

PERCEPTIONS OF TEACHER EVALUATION MODELS

CHAPTER TWO: LITERATURE REVIEW

This review of literature is separated into five different sections. The first section is an overview of the federal and state accountability era in education and the shifts in teacher evaluations. In the second section, I provide a historical account of teacher evaluation policies and practices. In the third section, I review changes to current practices and overall evaluation thinking and the various components to teacher evaluations. In the fourth section, I summarize the three different evaluation models (RISE, TAP, and PAR) and the similarities and differences found within each. My literature review then finishes with the fifth section where I am overviewing frameworks by Danielson and Marzano.

Federal and State Accountability Era in Education

Over the last several years a heavily-debated topic for those in the educational system in both the United States and abroad has been how teachers are evaluated (Darling-Hammond, 2013; Marzano & Toth, 2013). Policies and programs enacted in recent years have led to significant redesigns of the teacher evaluation process. Many of these changes to the teacher evaluation structure have occurred in large part because of federal programs and mandates such as NCLB (2002) and RTTT (Civic Impulse, 2016). The federal NCLB legislation and state accountability systems have placed greater obligation on principals and teachers to implement better evaluation protocols and procedures to increase student learning (Leithwood, Louis, Anderson, & Wahlstrom, 2004). NCLB focused on holding key participants, both the evaluator and evaluatee, accountable for student academic success.

The federal RTTT legislation resulted in dollars for selected school districts/states willing to mandate new, more effective teacher evaluation systems as one of the six priorities designed to help states reform their current overall educational system (Johnson & Stephens, 2012). The

PERCEPTIONS OF TEACHER EVALUATION MODELS

result of this potential financial lure of federal dollars really raised the ante for states to implement new, more accountable, and innovative “emergent” teacher evaluations systems (USDOE, 2012). In fact, since 2008, more than 30 states have enacted legislation that changed evaluations methods (National Governors Association, 2011). If states did not raise the stakes put on teacher evaluations, they automatically would not qualify for this federal funding. Given this fact, of all the educational practices today, none has been more examined than the way teacher evaluations are utilized and administered (Scannella & McCarthy, 2014). Therefore, an attempt to study the effects of these changes regarding teaching evaluations is extremely timely and necessary.

During this time of overall heightened school district educational accountability and the sharing of individual school performance data, the outcomes for teacher evaluations have never been more scrutinized. This has led to teacher evaluations being shared publicly for the first time by school and category. Two pieces of legislation were the forerunners to the Every Student Succeeds Act (ESSA, 2015): the RTTT legislation and the reauthorization of the Elementary and Secondary Education Act. The ESSA forced statewide and local assessment results to directly link to the evaluation of teacher quality and ratings. In Indiana and many other states, school districts historically were allowed to conduct teacher evaluations with minimal direction and guidance from the state.

During the last decade, there has been considerable federal and state attention, focus, and money spent on developing evaluation criteria and models to measure the performance of professional practices (Baker et al., 2013; Steinberg & Donaldson, 2016). The goals behind these efforts are to attract and retain effective teachers, improve teacher performance, and increase student achievement (Morgan, Hodge, Trepinski, & Anderson, 2014; Steinberg &

PERCEPTIONS OF TEACHER EVALUATION MODELS

Donaldson, 2016). The reasons for increased scrutiny and focus can be directly linked to criticisms of past teacher evaluations systems and practices (Baker et al., 2013; Cohen & Goldhaber, 2016; Ford, Van Sickle, & Fazio-Brunson, 2016). For example, in many evaluation models, the criteria for evaluations are vague in nature and focus more on results than on processes; in essence, more emphasis places on measures of student achievement rather than on teacher growth (Danielson, 2012; Marzano, 2012; Stonge, Tucker, & Hindman, 2004). My study will seek to analyze what consistent practices for the teacher evaluation models in the state of Indiana based on administrative and teacher perspectives.

Because of this, the teacher evaluation process has become a heavily-debated topic for those in the educational system in the United States (Knoeppel & Reinhart, 2008; Ravitch, 2010; Timperly, 2011; Zhao, 2009). The teacher evaluation process has always been done a little differently from state to state, even school district to school district. With very little teacher evaluation consistency to speak of, what has developed with evaluations has reached a level of critical importance (Knoeppel & Reinhart, 2008; Ravitch, 2010; Timperly, 2011; Zhao, 2009). Classroom education in general has long been at the center of the blame for poor student performance; thus, there has been a call for greater accountability for teacher evaluations (Gutherie & Springer, 2004).

History of Teacher Evaluation Policies and Practices

After looking at current legislation and mandates with teacher evaluations over the last thirty years or so, I will now review the evolving history of teacher evaluations. In order to foster an understanding of the issues connected with identifying the best teaching methods and evaluating the implementation of these methods, I will analyze the progression of teacher evaluation policies in the United States from its origin to current practices.

PERCEPTIONS OF TEACHER EVALUATION MODELS

In the 1700s, the field of education was not an area of study, nor was it really a professional discipline. The local government and clergy hired teachers and ultimately made the judgments about teacher success or lack thereof. Because a teacher was a community servant, there was no larger scale uniformity regarding the importance of pedagogical expertise, the need for feedback, or the methodology of providing that feedback (Danielson & McGreal, 2000; Darling-Hammond, 1996; Gitlin & Smyth, 1989). Over the next few decades, with the United States growing as an industrial nation, the need for education and the need to have common schools increased in larger, urban areas. In these urban schools, the expectation developed that teachers needed to be better qualified in specific disciplines (Natriello, 1990). By the mid-1800s, the nature of teaching was viewed as a more complex practice and the focus became centered on improving instruction (Ellett & Teddlie, 2003). These first 150 years saw tremendous growth in the awareness that teacher pedagogy plays a critical role in teacher effectiveness (Shinkfield & Stufflebeam, 2012). However, it was not until the 20th century when educators began to understand the value of and to implement teacher evaluations in a more formalized and meaningful manner (Darling-Hammond, 1996; Kellaghan & Stufflebeam, 2012).

During the early 20th century, teacher evaluations were widely administered only for the sake of completing a process of paperwork, not for the functional purpose of providing feedback to enable the teacher to improve. (Marzano, 2012; Protsik, 1995). Often the result was to just get the evaluations completed regardless of whether or not the teachers would be given information to improve on their instructional strategies (Darling-Hammond, 1996; Protsik, 1995). The two predominant and competing views of education were developed by John Dewey and Frederick Taylor. Dewey felt that schools should be a place where students practice citizenship and further develop the ideals of democracy (1938). Many current practices found their genesis in Dewey's

PERCEPTIONS OF TEACHER EVALUATION MODELS

thinking, such as differentiation to meet the student's learning needs and the modeling of engagement strategies by teachers for students (Danielson, 1996; Hall, 2002; Peterson, 2000; Tomlinson, 2014). Dewey's focus was on the ultimate end goal of education. In Dewey's words, "growth, not only physically but intellectually and morally, is one exemplification of the principle of continuity for educational learning" (1938, p. 5). From a much different viewpoint, Taylor believed that schools should be aligned with a business/factory model and viewed measurement as the best tool for a more scientific approach to schooling. Taylor's (2004) work was some of the first with regard to thinking of reliability, predictability, and logic, which connects with the bureaucratic classical theory. Under Taylorism, management assumed more duties by reducing teachers' responsibilities and decision-making, which led to the standardization of educational methods and processes (Marzano, 2012; Spring, 2010).

In his book, *Public School Administration*, Cubberley (1929) set out principles emphasizing that measurement and analysis of data could be used by school administrators to ensure teachers' productivity. Cubberley also provided specific examples of how the scientific methods could be used when evaluating teachers. For example, it is the business of the school to build its products (children) to the specifications laid out for schools. From the two competing views of Dewey and Taylor, many further ideas were expanded and developed in the 1930s – 1950s. During this time period, William Wetzell furthered the thinking on evaluations and distanced himself from "schools as factories" thinking (Wetzell, 1932). He recommended two components regarding teacher evaluation: the establishment of clear and measurable objectives for each course and the use of reliable measures of student learning. Teacher ratings and checklists evolved through the ensuing historical periods which drove the beginning thinking for the evaluation tools used in current times (Duffy, 1998).

PERCEPTIONS OF TEACHER EVALUATION MODELS

After World War II, the thinking concerning teacher evaluation began to shift away from Taylor's scientific raw materials and products approach to one that placed emphasis on the teacher as an individual (Danielson & McGreal, 2000). This shift increased the administrator's responsibilities and, without a corresponding development of evaluation methodologies, left the evaluation process in flux (Lagemann, 2002). During this 10-20-year period, conventional wisdom established the importance and utility of observing and providing feedback to teachers, and this understanding set the foundation for the movement to a clinical supervision approach (Lagemann, 2002; Newman, 1992).

Clinical supervision gained a stronghold in education in the early 1960s. The process involved a purposeful educational relationship between the teacher and evaluator, where observation and discussion were the measuring sticks for higher levels of growth and effectiveness (Anderson, 1993; Danielson & McGreal, 2000; Nolan & Hoover, 2011). The model that emerged from this work was published in Goldhammer's (1969) work, *Clinical Supervision: Special Methods for the Supervision of Teachers*. Goldhammer developed a five-phase process designed to involve the administrator-evaluator and teacher in reflective practice dialogue. The five phases were as follows: a) pre-observation conference, b) classroom observation, c) analysis, d) a supervision conference, and e) analysis of the analysis (Goldhammer, 1969; Neville & Garman, 1998; Smith & Andrews, 1989).

In 1972, Cogan wrote *Clinical Supervision*. He believed that evaluators must look for critical areas that could be impeding student learning and engage teachers in positive conversation regarding those areas. Cogan concluded that an evaluator's personal beliefs about good teaching can at times construct a roadblock to providing effective feedback:

PERCEPTIONS OF TEACHER EVALUATION MODELS

Most teachers have consciously or unconsciously constructed a personal model of a good teacher. Such conceptions generally grow by accumulation of knowledge rather than by critical examination and careful testing. The result is that too often the operating model of the teacher-turned-supervisor is pretty much what he himself does well. When teachers become supervisors, these personal preferences generally operate in full vigor, furnishing many of the criteria for viewing the teaching of others. (p. 54)

His words are as true today as they were some 40 years ago. The concepts clinical supervision introduced regarding the value of rich dialogue occurring between teacher and evaluator still drives ideas and discussions (Acheson & Gall, 2003; Anderson, 1993; Sullivan & Glanz, 2005; Zepeda, 2014). Concrete steps were being put into place establishing the acceptance and use of formal evaluations. However, it was not yet realized that the trusting dialogue that was taking place often times mirrored a ritualistic set of steps to be followed and completed in the quickest way. Another difficulty for the clinical supervision advocates was a reluctance to define any characteristics of what effective teaching and instruction might look like (Acheson & Gall, 2003; Cogan, 1972, Nolan & Hoover, 2011; Sullivan & Glanz, 2005).

This reluctance was countered, and the methodology of teacher evaluation again advanced, with the development of the Madeline Hunter Model of Lesson Design. In 1982, Hunter developed highly structured plans using the classic repetitive lesson model. The traditional steps of the Hunter model were designed for the explicit purpose of having students get it right the first time through (Hunter, 2004). She included in her teacher lesson plan template seven different resources and materials needed. The model included anticipatory set, objective/purpose, instruction input, teacher modeling, checking for understanding, guided practice, and independent practice (Hunter, 1982; Watts 2016). Ironically, Hunter was emphatic

PERCEPTIONS OF TEACHER EVALUATION MODELS

that it was never the intention that her model should be used as a teacher evaluation model.

Given that Hunter's model was centered on lesson planning, it did not capture the essence of the overall responsibilities and total expectations needed with comprehensive teacher evaluations (Danielson & McGreal, 2000; Hunter 2004).

For instance, Hunter knew that the human brain lays down pathways as it learns. Thus, she wanted to assure that teachers reduce learners' chances of "getting it wrong" so that incorrect neural pathways were not established. Hunter advocated this approach because the research at the time indicated that relearning materials or skills took much more time than learning it correctly the first time (Hunter, 2004; Stallings, 1985; Watts, 2016). However, it is clear that current evaluation practices are often closely aligned with her ideas (Marzano, 2012; Peterson, 2000; Saultz & Saultz, 2017; Taylor & Tyler, 2012; Tomlinson, 2014).

Teacher evaluation was once again at the forefront with the RAND study done during the early 1980s (Wise, 1984). The study was developed to discern what types of teacher evaluation practices were actually occurring in school districts throughout the United States (Darling-Hammond, 1996; Wise, 1984). Several components of the study produced valuable information moving forward in the teacher evaluation process (Marsh, 2007; Darling-Hammond, 1996; Stoelina, 2017). The study found that the reflective pieces were oftentimes too open-ended to improve teacher pedagogy. In fact, the study indicated that teachers wanted a more definitive and standardized evaluation process (Wise, 1984). Without a standardized process, teacher ratings and evaluations were inconsistent, and thus, the influence of the teacher unions in teacher evaluations began. This is consistent with what has happened in the Indiana school districts with models and expectations allowed to be a local district selection.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Another interesting point to the study was the survey feedback indicated a belief that administrators were not qualified to do the evaluations, thus, at times triggering teacher resistance to the feedback given. Finally, the study emphasized the lack of training for evaluators to provide teachers with useful classroom evaluations (Darling-Hammond 1996; Marsh, 2007; Wise, 1984). These findings helped bring the focus on several questions, including: How do teacher evaluations improve if the teacher does not want to self-reflect and have clear goals for improvement? How can teacher evaluations be a reflective process without being so punitive?

To answer these and similar questions, in 1996 Danielson published *Enhancing Professional Practice: A Framework for Teaching*. The framework was divided into 22 components clustered into using four domains of teaching responsibility (Danielson, 1996; Darling-Hammond, 2013). This success of the framework was based on the need to define high-quality teaching. Additionally, the framework set up clear and succinct definitions of teaching while being aware of the complexities that a teacher must handle (Danielson, 1996; Darling-Hammond, 2013; Tomlinson, 2014). Because this framework continues to be used as a reference point for any new models developed today, it is important to include it in the history of teacher evaluations. Also important to note is that the initial intention of Danielson's work was not to be used to evaluate teachers. It was created as a tool that would provide a structure to help teachers grow in their professional abilities yet not be scrutinized or evaluated (Danielson & McGreal, 2000). The intent was to facilitate teachers' growth and improvement in classroom duties without the pressure of being penalized during the process. Given the need for growth and evaluation to take place simultaneously, reformation of school processes needed to occur. For example, there had to be a redesign in the responsibilities, roles, and relationships between

PERCEPTIONS OF TEACHER EVALUATION MODELS

teachers and administrators (Danielson & McGreal, 2000; Darling-Hammond, 2013; Marzano & Toth, 2013). Since the reformations, evaluation systems that support teacher growth and place emphasis on teacher needs have taken Danielson's original work to better results (Danielson, 2012). The next important piece to look at would be the evolution of recent changes and requirements of teacher evaluation.

Research on Teacher Evaluations in Indiana

Since the change in the law, various studies have been done regarding teacher evaluations in Indiana. The studies have provided scholarly analysis of the changes over the last few years. However, in my research, I found there to be nine such studies regarding teacher evaluations in Indiana, and most of the studies (seven of nine) involved quantitative measures using surveys to gain feedback about evaluations in the context of the specific aspect being studied. For instance, Boyland, Harvey, Quick, & Choi (2014) examined information regarding the then-existing status of teacher evaluations in Indiana. In that study, 477 elementary and secondary principals were surveyed and asked about their use of summative and formative evaluation models in order to provide a baseline understanding of the then-current models. The purpose of the study was to determine existing perceptions toward evaluations so that the future changes being mandated by the new laws could be accurately measured. The study revealed that there was a low use of the new evaluation methods and that principals disagreed that the new methods would be effective.

A similar study was conducted in Indiana to investigate whether evaluation practices that foster principal effectiveness are important to improving professional practices that increase educational opportunities for all students (Andrews, 2015). This study investigated and compared the perspectives of superintendents and principals with regard to the RISE Principal Evaluation and Development System utilized to evaluate principals in Indiana. Survey

PERCEPTIONS OF TEACHER EVALUATION MODELS

methodology was employed to gather feedback from evaluators (superintendents) and those being evaluated (principals). In this study, a total of 364 school leaders participated, and by using that data, the study measured the differences between principals' and superintendents' perceptions of model efficacy.

Another study conducted in Indiana identified whether a relationship existed between the implementation of professional evaluation processes and the use of research-based teaching practices, factoring in both perceptions of principals and practicing teachers (Sargent, 2014). The variables of professional development on the evaluation model and the principal's years of experience, degrees contained, and types of degrees were factored into the analysis. For this quantitative study, principals were surveyed to identify the teacher evaluation model used in the school along with professional development, years of experience, degrees, and types of degrees. In addition, the principals identified the use of research-based teaching practices in the school, prior to and after implementation of the teacher evaluation model. Teachers within the evaluation model were surveyed to ascertain the use of research-based teaching practices, prior to and after implementation of the model within their schools. The Sargent study determined a relationship existed between principals' and teachers' perceived use of research-based teaching practices after the implementation of the teacher evaluation model.

In a qualitative study done by Eberline (2016), physical educators were surveyed to determine whether being evaluated under the required, new models caused a difference in teacher effectiveness. Eberline interviewed 22 teachers in 15 school districts in Indiana. Teachers face a difficult task of using limited time and resources to fully impact students. And, since physical education is labeled as a noncore subject with no standardized test to evaluate student learning, it is hard to measure growth by teachers and students. The purpose of the study

PERCEPTIONS OF TEACHER EVALUATION MODELS

was to describe physical educators' perceptions of and experiences with the teacher evaluation system in the state of Indiana. Additionally, this research examined challenges faced by teachers as they adapted to teacher evaluation systems. Solutions to the evaluative mandates were sought to address the shortcomings of Indiana teacher evaluations in physical education.

Given that the vast majority of studies conducted in Indiana have involved quantitative research (seven of nine studies), my study will contribute to the body of qualitative research on teacher evaluations in Indiana post-P.L. 90 by providing insight into the administrator and teacher perceptions being studied.

Changes to Current Practices and Standards-Based Evaluations

Beginning as early as 2007, more than two-thirds of the country's states had passed legislation that called for new methods in evaluating teachers (Koppich & Humphrey, 2011; Mathis, 2011). As states have passed legislation mandating teacher evaluation reform, local districts have been called upon to implement new systems for evaluating the effectiveness of their teachers. Traditionally, in many local school districts, evaluation systems were subject to collective bargaining and, therefore, highly influenced by teacher unions (Spring, 2010). Several states have removed teacher evaluation from the scope of collective bargaining to give district administrators the freedom to implement expensive teacher evaluation reforms (Shaha et al., 2015). Besides the potential loss of federal dollars, the reason for the states' increased scrutiny centered on the fact the evaluation process being done differently from state to state and even school district to school district. Given these differences, how to measure the evidence of teacher contributions to student learning has become a much greater conversational element concerning value-added components to teacher evaluation models (Darling-Hammond, 2015; Tomlinson, 2014). The push to make teacher evaluations more outcome-driven also increased

PERCEPTIONS OF TEACHER EVALUATION MODELS

the need to look at the various models that were not viewed as providing useful and timely feedback to help teachers grow professionally (Goldring, Grissom, Rubin, Neumerski, Cannata, Drake, & Schuermann, 2015). Yet, it is not surprising that some states and school districts have ignored some of these crucial components (DiCarlo, 2012). Components were often ignored because analysis of relationships between professional practices and student achievement gains are hard to measure from a teacher effectiveness perspective (Taylor & Tyler, 2012).

In fact, before the required changes were implemented, evaluations usually consisted of observation checklists that included competencies for a building leader to “check off” to rate teacher performance (Munson, 1998). These checklists did not capture the details of what took place in the classroom; therefore, little evidence could be found to support an authentic determination of effective teachers (Danielson, 2011; Protheroe, 2002). In fact, the history of the evaluation system shows ineffective methods or tools have been used and further professional development has not been tied to evaluations (Danielson, 2011; Toch & Rothman, 2008). As schools and districts across the country are requiring measurable improvements in student achievement, it is important that high-quality teaching and high-quality teachers be identified. To be open to finding the effects of high-quality teaching, a valid and reliable method of identifying quality instruction is necessary. Yet previous research has shown that traditional principal evaluations of teachers are inadequate both for differentiating between more and less proficient teachers and as a basis for guiding improvements in teaching skills (Gallagher, 2004). The end result was nearly 98% of teachers receiving a “satisfactory” designation (Schachter, 2012).

The desire for consistency was instrumental in evaluations reaching a mandated status towards greater accountability and uniformity regarding teacher evaluations in both overall scope

PERCEPTIONS OF TEACHER EVALUATION MODELS

and sequence. A long held educational belief is that classroom practices and teacher quality in general have consistently been at the center of the blame for poor student performance (Looney, 2011). The call for increased accountability forced states to pass new legislation requiring a) teachers to undergo regular evaluations at least once a year and b) administrators to increase the rigor of the individual teacher evaluations so that they would be conducted systematically and with fidelity (Looney, 2011). In fact, for current practice to advance and for students to compete in the global economy, there must be an evaluation system with a focus on effective instruction and annual student growth (Oliva et al., 2009).

Current Best Practices Regarding Teacher Evaluations

Some commonalities with regard to the current best practices regarding teacher evaluations follow.

Overall framework

To meet the government-imposed expectations, new teacher evaluation models were developed and used in schools for decision-making regarding employment, compensation, and teacher growth. The new models purport to achieve the desired consistencies; however, many educators question the validity of those claims, pointing specifically to the lack of common definitions, standards, and enforcement procedures (Danielson, 2011; Darling-Hammond et al., 2012). In the end, these and other questions should be answered to focus on the twin goals of a) ensuring teacher quality and b) promoting professional development (Danielson, 2011). The ultimate goal of teachers positively impacting students' abilities to demonstrate learning and knowledge can be furthered by an effective (i.e., fair and consistent) teacher evaluation process.

Expectation of quality

Two elements at the core of educational practice are teaching and learning, and the research is clear that teacher quality is the most important school-level factor affecting student

PERCEPTIONS OF TEACHER EVALUATION MODELS

learning (Dillon, 2013; Looney, 2011; Marzano & Toth, 2013; Selwyn, 2007). In fact, the method of evaluation may not only affect which specific teachers are rewarded in the short term, but also shape the qualities of teachers and the teaching students experience in the long term (Harris, Ingle, & Rutledge, 2014). Research has been done to understand what characteristics and components are needed to have an effective teacher evaluation. A teacher evaluation should provide teachers with useful feedback on classroom needs, the opportunity to learn new teaching techniques, information on how to grow professionally, data points, and ongoing professional conversations (Danielson, 2011; Darling-Hammond, 2013; Marzano, 2012). Using feedback in teacher evaluations is a significant topic in educational circles because it allows the evaluation process to improve the quality of teaching instead of being utilized solely for accountability purposes (Danielson, 1996; Danielson & McGreal, 2000; Goe, 2013; Natriello, 1990).

Problems to be overcome

The new teacher evaluation structure implemented in education presents a complex set of problems, and two particularly important ideas are at the forefront of the discussion. First, research indicates that teachers are the most powerful school-based factor affecting student achievement (Hanushek & Rivkin, 2010; Kane & Staiger, 2008), yet teachers tend to have the perception that evaluations are only punitive in nature; they do not view them as instruments of professional growth. Second, effectiveness varies greatly between teachers, and those differences in effectiveness are not well predicted by traditional evaluation models, especially considering the lack of consistency among administrators in implementing and interpreting the evaluation tools. “If the goal of evaluations is to grow great teachers to drive student excellence, the traditional model has failed” (Bambrick-Santoyo, 2012, p. 28). To address these problems, systems such as Danielson’s (1996) *Framework for Teaching*, have provided educators with a

PERCEPTIONS OF TEACHER EVALUATION MODELS

model for teacher performance that encompasses research-based practices for effective instruction, but the use of these systems is not consistent nor widespread.

In addition to these two philosophically-based problems, there exists the universal obstacle of time. “Time is the greatest obstacle to making teacher evaluations as useful as they need to be” (Zatynski, 2012, p. 24). Conducting quality evaluations and providing relevant, useful feedback for teachers takes time, and evaluators who are engaged in the day-to-day operations and procedures of the buildings they manage can find it difficult, if not impossible, to carve out the hours in a day needed to accomplish the task. Yet, high quality feedback is imperative to the level of success of the evaluations. Similarly, money can be a considerable factor school districts must weigh if the evaluation model they have might require the heavy use of trainers, teacher personnel, or outside consultants.

Competing views

Given the desire to increase student achievement, the need to improve teaching and learning, and mandates by the federal government that are reflected in requirements for grant programs such as RTTT and the Teacher Incentive Fund, many states have developed teacher evaluation systems that include measures of individual teachers’ contributions to their students’ learning growth. Although these recent federal mandates requiring the use of new teacher evaluation systems conflict with the way things have been done in the past, to some it was a necessary step to bring about needed change in education. Sartain, Stoelinga, and Brown (2011) specifically pointed out that traditional evaluation systems failed to give meaningful feedback on instruction and were not differentiating amongst effective and ineffective teachers. In other words, previous evaluation systems were inadequate. Often times, under the banner of teacher autonomy, teachers would be left more or less to their own devices (Darling-Hammond, 2013),

PERCEPTIONS OF TEACHER EVALUATION MODELS

and the faulty systems identified nearly all teachers simply as satisfactory or not satisfactory, without offering any useful feedback or direction to teachers on how they can improve. With that said, it is easy to find others who believe that those schools that are chasing the federal funding consider education as a bottom-line endeavor and actually just do not trust or respect teachers as true professionals (Marzano, 2012). Given both sides of thought, the development of an evaluation system that genuinely encourages professional growth, improved classroom practice, and increased teacher feedback continues to be at the center of focus. The need for increased teacher feedback cannot be overstated. Feedback is one of the most powerful influences on learning and achievement by teachers, but the impact of feedback can also be positive or negative if not done properly (Hattie & Timperly, 2007).

Teacher evaluations as a growth model

Teacher evaluations must support continuous growth and improvement, both for individual teachers and for the teaching profession itself (Darling-Hammond, 2013). For decades, teacher evaluations were more a bureaucratic exercise that failed to recognize either mediocrity or excellence in teaching. As such, best done teacher evaluations represented an enormous opportunity for giving teachers valuable feedback that could help their practice.

“Measuring a teacher’s impact on student learning lets a teacher and her supervisor know whether what she is doing in her classroom day to day is working” (Almy, 2011, p. 1). Also,

[s]ystems that help teachers improve and that support timely and efficient personnel have more than good instruments. Successful systems use multiple classroom observations across the year by expert evaluators looking at multiple sources of data, and they provide timely and meaningful feedback to the teacher. (Darling-Hammond, 2012, p. 13)

PERCEPTIONS OF TEACHER EVALUATION MODELS

These are just some of the multi-faceted needs of a great teacher evaluation. Studies show that the teacher evaluation system is in need of some updates and changes. Danielson (2011) addressed the need for a teacher evaluation system to be developed that would better support conditions for increasing both teacher and student learning. An evaluation should be a process that improves teaching in a classroom. Danielson also outlined areas of shortcomings with current evaluation systems in poor evaluation criteria, ambiguous language, and inconsistency in meeting different types of teacher needs, lack of evaluator uniformity, and a process that is very bureaucratic in nature. An evaluation model that creates opportunity for growth in teachers, ensures teacher quality, and guarantees credibility in the educational field would need to be purposefully designed to address these shortfalls. This study will help to determine what current parts of researched models might meet these addressed areas of requirement and what areas still need to be improved to meet the theoretical best practice evaluation models. In fact, the key driver for teacher development is the guidance on exactly how to improve (Bambrick-Santoyo, 2012). In order to do this, there must be a consistent definition of expectations that has a shared understanding by everyone.

Once a clear vision is established, the next step is to have trained evaluators who can make fair, reliable, and valid programmatic and personnel evaluations. Danielson (2011) stated that another purpose of the teacher evaluation is to promote professional learning because the science of teaching is so difficult that educators, both administrators and teachers, should always look for ways to improve it. Unfortunately, educators' attempts at merging quality assurance with professional learning had varying degrees of success. In this research, Danielson (2011) discussed a two-year pilot study that reinforces her ideas. In the study, three areas were found to be critical to helping teachers learn. First, a common definition by evaluators is the critical

PERCEPTIONS OF TEACHER EVALUATION MODELS

beginning piece. Second, more opportunities to conduct meaningful observations and engage in professional conversations must take place. Third, there must be a clear definition of what matters in the teaching and learning of students. According to Danielson (2011), the ability to organize these three critical pieces becomes the cornerstone for evaluations that help teachers learn and grow in their educational practice.

Other key components of teacher evaluations

Koppich and Humphrey (2011) also addressed the pitfalls of the current system and outlined the need to get serious about teacher evaluation through the development of peer assistance and review. The premise is that a new teacher evaluation system is critical. It must be intentional, focused, comprehensive, and rigorous in scope (Burris & Welner, 2011; Danielson, 2011; Ellett & Teddlie, 2003). A study was conducted by the authors in two different districts that focused on peer assistance and review. According to studies, peer assistance and review occurs when carefully selected consulting teachers work with beginning and underperforming teachers (Goldstein, 2008; Keller, 2006; Papay & Johnson, 2012). There was a pre-arranged agreement with the superintendent, school board, and teacher association to pay these consulting teachers for their involvement in the evaluation process (Koppich & Humphrey, 2011; Munson 1998).

In the studies, one area really stood out: the integration of support of fellow mentor teachers, department leaders, and administration with the evaluation yielding considerable teacher learning and growth (Johnson & Fiarman, 2012; Munson, 1998). The consulting teachers completed comprehensive evaluations, which included the documentation of sometimes up to 150 pages of detailed notes and information.

PERCEPTIONS OF TEACHER EVALUATION MODELS

The authors discovered through the study that principals do not have the time to support and conduct the evaluation reports for the teachers to this magnitude. The practice of observing classrooms and using learning walks is an example of moving to professional practice by school staffs (Danielson, 2012; Fisher & Frey, 2014; Johnson & Fiarman, 2012). With the evidence the consulting teachers were able to amass, the teacher ratings were based less on emotional factors and driven more by data. In fact, a PAR panel often times requires consulting teachers to substantiate all evaluations and observations with specific evidence grounded on the standards (Johnson & Fiarman, 2012; Kane & Cantrell, 2012; Keller, 2006).

In Koppich and Humphrey's (2011) comprehensive teacher research, conversations amongst evaluators and evaluates were recorded that focused on intensive questioning and probing regarding teaching and learning in the classroom. In these type of studies, governance boards were created to have oversight of the evaluation process (Keller, 2006; Koppich & Humphrey, 2011; Munson, 1998). Their job was to make sure no one short-changed the process and that authentic evaluations were completed. If any group was lacking, the person(s) would be sent back to redo their work. Overall, what researchers found was that this process of teachers and administrators working together to strengthen teacher evaluation not only helped improve previous union relations, but also enabled teachers to learn and improve at a much faster growth pace (Danielson & McGreal, 2000; Johnson & Fiarman, 2012; Keller, 2006; Marzano & Toth, 2013; Sartain et al., 2011). Like others who have studied peer assistance review (Goldstein, 2008; Papay & Johnson, 2012), it was found that districts that fully implemented programs had better teacher retention rates.

In summary, new evaluation tools, notably improvements in classroom observations and the ability to link teachers to student achievement, have been developed that are more accurate

PERCEPTIONS OF TEACHER EVALUATION MODELS

measures of teacher and school performance. Much of the research being reported centers on the importance of administrators working in partnership to select a standards-based teacher evaluation model (Danielson & McGreal, 2000; Darling-Hammond, 2013; Taylor & Tyler, 2012; Tomlinson, 2014; Zeichner, 2011). There is a need for a wide range in training for both administrators and teachers so that a common understanding of effective teaching can be reached (Danielson & McGreal, 2000; Darling-Hammond, 2013; Taylor & Tyler, 2012; Tomlinson, 2014). Finally, the need exists to bring the teacher evaluation system in line with the overall goal of creating a quality instructional program.

Overview of the Three Evaluation Models (RISE, TAP, and PAR)

Because of the legislative mandate in the state of Indiana in 2011, a new process for evaluations was put into action. Based on this mandate, educational organizers needed to develop benchmarks for credible and accurate evaluations that measure teacher performance and professional growth. The State allowed the use of any teacher evaluation model as long as it met three criteria. First, the evaluation must at least be completed annually; every teacher, regardless of experience, warrants meaningful feedback on their performance on an annual basis. Second, the evaluation must include student growth data; evaluations should be student-focused. An effective teacher helps students make academic progress. A thorough evaluation system includes multiple measures of teacher performance, and growth data must be one of the key measures.

Third, the evaluation must include four rating categories; to retain the best teachers, a process is needed to differentiate the best educators and give them the recognition as such. If teachers are to perform at the highest level, administrators need to know which individuals are achieving the highest success and which need support because they are new or having difficulty

PERCEPTIONS OF TEACHER EVALUATION MODELS

in teaching students. Modeled after these criteria, teacher evaluation models were constructed and implemented across the state of Indiana.

Background on three specific models most used by school districts in Indiana follows. When the new law came into effect, these three models were the most talked about in terms of utilization, and that trend has continued. However, over the course of the five-year time period, the modified RISE model has had a highest trend of usage by districts to meet teacher evaluation expectations.

RISE

Educators think that teacher evaluations should be developed and implemented for the purpose of professional growth and improvement of teaching (Danielson & McGreal, 2000). Using this key premise, the Indiana Department of Education (IDOE) began to research different models to evaluate teachers. The RISE evaluation and development system (RISE) was a model developed over the course of a year by the Indiana Teacher Evaluation Cabinet, a diverse group of educators from around the state, more than half of whom have won awards for excellence in teaching. This work was circulated widely to make sure that efforts represented the best thinking from Indiana educators. The Cabinet continued to refine RISE based on feedback from the principals and teachers who used it every day. The RISE model relies on multiple sources of information and is thought to paint a fair, accurate, and comprehensive portrait of a teacher's performance, and its components closely resemble the elements found in Danielson's (1996) *Framework for Teaching*.

Through classroom observations and conferences, RISE provides a clear picture of what teachers do in their classrooms and schools. RISE provides information on the most important aspects of teaching: planning, instruction, leadership, and student growth. RISE identifies

PERCEPTIONS OF TEACHER EVALUATION MODELS

teachers' strengths and development needs and also recognizes excellent teachers and encourages them to share their best practices (IDOE, 2012). But more importantly, RISE encourages *all* teachers to improve their instruction and grow as professionals. RISE sets high expectations for principals as well, encouraging them to collaborate with teachers around a shared vision of quality instruction (IDOE, 2011).

TAP

The model called Teacher Advancement Program (TAP) is a comprehensive school reform with its stated goal to attract, retain, and motivate quality teachers. By aggressively recruiting new teachers, providing a career continuum, introducing teacher-led professional development, implementing rigorous teacher accountability, and paying teachers based on their position, teaching skills, and how much their students achieve, TAP schools change their organizational structure to support and reward high-quality instruction (Solmon, White, Cohen, & Woo, 2007). Under the TAP model, teachers are evaluated at least three times a year by teams of "master" teachers who are trained to use the TAP evaluation rubric. This rubric assesses four domains entitled Instruction, Planning, Environment, and Professionalism, and each domain has its own independent rubric (Pieczena, 2012).

PAR

The model called Peer Assistance and Review (PAR) involves teachers measuring the accountability of other teachers. PAR levels the traditional institutional hierarchy by allowing teacher peers to observe, compared to only principals and other administrator-level educators in other models. More specifically, in a peer review, a consulting teacher (CT) guides and supports a participating classroom teacher based on a professional development need. Often times, CTs are no longer classroom teachers and report to a panel made up of teachers and administrators.

PERCEPTIONS OF TEACHER EVALUATION MODELS

These carefully chosen, experienced teachers provide intense support to beginning teachers and then conduct their evaluations. In many cases, if a teacher is found to need support or does not meet evaluation standards, then he or she is assigned a CT. Consulting teachers diagnose PAR teacher's strengths and weaknesses and offer tailored one-on-one support, and in doing so, they "need to know how to build trust with colleagues while maintaining high expectations" (Johnson & Fiarman, 2012, p. 21).

In the PAR model, instead of dismissing an ineffective teacher, support is provided. The goal is not to get rid of ineffective teachers but instead to support them in the process of developing pedagogy (Keller, 2006). This evaluation model follows the belief that that whoever is in charge of improving teachers can, in fact, also be in charge of evaluating them. When teachers are allowed to control the quality of teaching, through evaluation and professional development and professional development, they are vested with the power to be open to finding who is a good teacher and who is not (Goldstein, 2008).

Similarities and Differences among the Three Models

In Table 1, I have highlighted some of the central principles from each of the three models of observation. In all three models, observations are conducted and completed by three different groups of evaluators, which is significant in the scope of effectiveness. The second main difference is in the all-important category of feedback and intervention. RISE appears to be the model that provides the most prescriptive evaluation setup with more tasks completed than actual feedback utilized. The TAP and Par models really focus on the importance of feedback and how it can inform decisions of next steps in the evaluation process. In the TAP and PAR model this feedback also leads to the formation of the desired course of action for professional

PERCEPTIONS OF TEACHER EVALUATION MODELS

development. Table 1 captures other similarities and differences between the RISE, TAP, and PAR models.

In summarizing the chart, all three models aimed to improve teacher effectiveness, but have different components and expectations to be followed. For example, there are differences in the number of categories that a teacher can be placed, differences in whom does the actual evaluations and in the frequency of evaluations, and differences in techniques to ensure teachers are being evaluated fairly and accurately.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Table 1

Similarities and Differences with the Three Models

	RISE Evaluation and Development System (RISE)	Peer Assistance and Review (PAR)	Teacher Advancement Program (TAP)
Categories of Teacher Performance	Four: Highly Effective, Effective, Improvement Necessary, and Ineffective.	Two: Satisfactory and Unsatisfactory.	None.
Measures of Input	Three domain rubric with an additional evaluation of "Core Professionalism".	Progress on Teacher-authored goals. Checklist based on the Toledo Plan.	Rubric based on the Danielson framework.
Measures of Output	Indiana Growth Model data and locally developed and assessed Student Learning Outcomes.	None	Teacher value-added and School value-added.
Observations	Conducted by team of trained primary and secondary evaluators.	Conducted by peer teachers and principals.	Conducted by mentor teachers, master teachers, and principals.
Teacher Feedback and Interventions	Optional Mid-Year Check-in. Loosely structured Professional Development Plan. Professional Development Plan required for Ineffective or Improvement Necessary.	Teacher and Consulting Teacher collaborate on teacher goal-setting. Teachers or principals can initiate an intervention prescribed by the plan.	Professional development informed by observations and performance data.
Summative Evaluation Determination	Three formulas for combining measures of input with measures of output. Formula be open to finding d by availability of IGM data for teachers' course loads.	Teachers must not receive any Unsatisfactory ratings on the final evaluation checklist.	Not explicitly be open to finding d by TAP. Corporations can use TAP data to make these determinations themselves, though.
Compensation Plan	No.	No.	Yes.

PERCEPTIONS OF TEACHER EVALUATION MODELS

The fact that federal programs have driven states to revise their teacher evaluation procedures is well established in this literature review. In order to qualify for the financial incentives, states were required to follow guidelines provided by the USDOE in 2012. At this point, I have briefly described each of the three evaluation models—RISE, TAP, and PAR—to provide the reader with a baseline understanding of each. However, before the three models can be compared to the Personnel Evaluation Utility Standards, it is necessary for the reader to understand an initial problem that compliance with the standards creates for the implementation of the models.

These guidelines clearly cross the threshold between formative and summative evaluations. There is a clear difference between the two types of evaluations. Formative evaluation was intended to be open to finding the merit of the educational program versus summative evaluation, which was intended to be open to finding the worth of the already finished educational program (Popham, 2013). Stated another way, formative teacher evaluation is focused on improvement, and summative teacher evaluation is focused on rewards or repercussions if minimal standards are not met (2013). The crux of the problem then is outlined by the fact that the guidelines that must be met include both formative and summative components.

In the review of the literature, it has been found that mixing both formative and summative evaluation expectations simply does not work. Formative components are used for growth in the process of the evaluation system while the summative part expects to see a final product, which then can be used in an accumulation of the evaluation as a whole. Therefore, as I began to compare the three evaluation models—RISE, TAP, and PAR—with the evaluation standards, I understood that all three models were already working from a point of view that

PERCEPTIONS OF TEACHER EVALUATION MODELS

teacher evaluations lean more heavily on subjective motives when thinking about human nature and teacher self-preservation behaviors.

Frameworks: Danielson and Marzano

According to Danielson, the two goals for teacher evaluation are professional development and quality assurance (Danielson, 2011). Danielson's framework seeks to provide detailed and objective criteria to describe teaching behaviors in evaluations. The intent is to reduce the subjectivity that is created by different experiences and perceptions among administrative evaluators. According to Danielson, the teacher evaluation systems have not accurately measured teacher quality and have not produced a highly proficient teacher workforce (Danielson & McGreal, 2000; Danielson, 2011). The model outlines the need for training of evaluators and a clearer understanding by teachers and administrators about what is great teaching (Danielson, 2011).

According to Marzano, evaluations need to be more than snapshots of instruction that take on measurements of ability, but fall short on context (Marzano, 2015). Marzano's framework works to develop a correlation between instructional strategies and student achievement. This model is organized into four domains: (a) classroom strategies and behaviors; (b) planning and preparing; (c) reflecting on teaching; and (d) collegiality and professionalism. The four domains contain 60 elements and collaboration tools for scoring that establish a knowledge base for developing expertise by teachers and administrators (2015). The rest of the Marzano's research focuses on an evaluation system that is comprehensive and specific, an evaluation system that includes a developmental scale, and an evaluation system that acknowledges and rewards growth. The Marzano model includes looking at all of the elements that research has identified as tied to student achievement.

PERCEPTIONS OF TEACHER EVALUATION MODELS

The pros and cons of the comprehensive models can be summarized into the following statement: No matter how many categories teachers are proficient in, if there is no true method in place to seek to improve their competence in the classroom, then it only measures and it does not change the teachers (Almy, 2011; Danielson, 2012; Darling-Hammond, 2015; Marzano, 2012; Nolan & Hoover, 2011; Peterson, 2000). The developmental model is a scale where teachers can track their skill development at levels from “not using” to “innovating.” Each of these different levels are designed to enable teachers to identify current levels and to set goals for improvement. In the rewards and growth model, teachers identify elements they wish to improve and then monitor progress all year. The teacher sets specific targets to reach, and then at the end of the year would receive an overall growth and status score. The overall findings suggest that an evaluation model that has both comprehensive and specific growth focus will in fact be the one that has the greater chance of better teacher quality and greater skilled teachers (Danielson, 2011; Marzano, 2015; Shaha et al., 2015; Tomlinson 2014).

In summarizing all the research, there is a definite need to change the past practice of teacher evaluation and move to a more effective model to improve overall teacher performance, which in turn would lead to higher levels of achievement by the students in schools (Alexander, 2016; Danielson & McGreal, 2000; Darling-Hammond, 2013; Marzano, 2012; Marzano & Toth, 2013; Tomlinson, 2014). A model must be developed to clearly place high expectations on professional growth and, thus, create a culture of evaluation in schools. The evaluation process described in the research is one that sees the balance between needing to know what goes on in the classroom and giving the teacher room to grow (Marshall, 1996; Marzano 2015; Peterson, 2000). Evaluation must increase the desire for teachers to engage in dialogue with supervisors and ensure that ongoing professional development be tied to goals. There must be work

PERCEPTIONS OF TEACHER EVALUATION MODELS

established to create clear and consistent definitions before any productive discourse can take place. To assess the quality of teaching practice, it is essential to define it. It is not sufficient to say, “I can’t define good teaching, but I know it when I see it” (Danielson, 2011, p. 36).

Summary

Interpreting the literature, the research uniformly points to the need for the development of a new and improved teacher evaluation model. More specifically, the new evaluation tool must accomplish what is known that effective evaluations should do—and that is help to improve teaching practices and increase professionalism in the teaching profession. From this, I believe it goes without saying, the major purpose of evaluation is to improve instruction and to develop the thinking of professionals who were involved in the process. In the research, steps were discussed for the setup of the teacher evaluation process. It requires getting teacher input and reaching a common vision or clarity for the final tool and the purpose behind the evaluations. Then, the process works to obtain common definitions, protocols, and language that can be put in place. The evaluation process should analyze multiple measures to compensate for any bias or flaws of each individual category to produce more effective evaluations (Kane & Cantrell, 2012; Tomlinson, 2014).

Next, the systematic process of putting together a method of professional development must quickly follow for all of the participants evaluated and the evaluators. Following that up, many pieces of literature describe putting in place a model for growth and learning that the teachers would find professionally rewarding. In every article and data source collected there is a common thread, look at peer teaching and coaching models/programs as positive factors for growth and change in the evaluation of teachers. The peer review model very clearly outlined the important role that consulting teachers can play in providing support to new and struggling

PERCEPTIONS OF TEACHER EVALUATION MODELS

teachers. The consulting teachers can offer this support and also conduct evaluations at the same time. For authentic feedback, the evaluators must be trained and skilled to ask the tough question and to get to the heart of the matter for real buy-in and change to develop.

As stated throughout the research, it is the professional conversations and dialogue that take place in the evaluation process where much of the growth occurs (Danielson, 2011; Darling-Hammond, 2013; Frontier & Mielke, 2016; Marzano, 2012; Timperley, 2011). I also find that teacher leadership is essential to productive teacher evaluations and ultimately supports growth in practice. Teachers grow their practices through effective professional learning and coaching. It was my goal to research teacher and administrator perceptions about the three main types of evaluation models currently used in Indiana. I intended for this study to enhance the thinking on current research and beliefs with these three teacher evaluation models.

PERCEPTIONS OF TEACHER EVALUATION MODELS

CHAPTER THREE: METHODOLOGY

This chapter describes the research design and methodology for this study. An explanation is provided including a rationale and assumptions for the qualitative design, the researcher's role, site and sample selections, managing and recording data, data collection techniques, data analysis procedures, and a summary of the chapter.

Rationale and Assumptions for the Qualitative Design

The reason for this qualitative research study is to understand teacher and administrator perceptions of completed evaluations when one of three High-Stakes Accountability Models—RISE Evaluation and Development System (RISE), Teacher Advancement Program (TAP), or Peer Assistance and Review (PAR)—is used. Data was collected on teacher and administrator perceptions concerning: (a) the general methodology or structure of the three teacher evaluation models; (b) the specific practices for gathering and compiling the raw data used in each of the three models (i.e., the forms used); and (c) the potential impact of each model on teacher effectiveness, student achievement, and professional growth. My qualitative research was based on these teacher and administrator perceptions and centered on naturalistic inquiry (Patton, 2002) with interviews and observations conducted in three different high school settings. Naturalistic inquiry involves research conducted without a predetermined or controlled setting or agenda and instead lets information gathering occur naturally and without manipulation (2002).

The design of this study linked the research questions to conclusions (Yin, 2013). This type of research focuses on in-depth understanding of social and human behavior and the reasons behind such behavior (Creswell, 2013; Hoy, 2010; Lichtman, 2013). Qualitative research is exploratory in nature, and therefore, I used this qualitative study to gain a rich description and understanding of teacher and administrator perceptions regarding teacher evaluations. This

PERCEPTIONS OF TEACHER EVALUATION MODELS

method of discovery, insight, and understanding from the perspective of those being studied was my focus because it provided insight into each school's teacher evaluation process story through the participant's stories (Marshall & Rossman, 2011; Merriam, 2009; Patton, 2002).

Qualitative research is preferable when the researcher needs to hear the stories and experiences of others (Creswell, 2013). It allows for consideration of the information gathered from the interview questions to provide insight into the differentiated perceptions of both administrators and teachers (Creswell, 2013; Patton, 2002). The meaning of this study and the reason for the qualitative nature was to gain the rich feedback and have the ability to see what is subtle but significant provided by educational professional practitioners (Eisner, 2017). This allows the reader a rich, in-depth view through the lens of administrators and teachers of varying experience levels to conceptualize the data from both the administrator/teacher emotional and intellectual levels. Advantages to using qualitative research for this educational study include very little hindrance by complex variables and not having to account for every variable in education (Bogdan & Biklen, 1998; Mears, 2009).

The rationale for using the case study design instead of other qualitative methodologies was to grasp a clearer representation of teacher evaluations within their natural setting. Additionally, the multiple case study approach results in the subject matter of the study being approached within its own context and from multiple sources of evidence, with data needing to converge in a triangular fashion (Baxter & Jack, 2008; Stake, 2005; Yin, 2013). As Baxter and Jack (2008) made clear, "this ensures that the issue is not explored through one lens, but rather a variety of lenses which allows for multiple facets of the phenomenon to be revealed and understood" (p. 544). According to Merriam (2009), the case study approach results in a rich and holistic account of teacher evaluations. It offers insights to expand its readers' experiences

PERCEPTIONS OF TEACHER EVALUATION MODELS

and has proven useful for studying educational programs and informing policy. For this multiple case study, I researched the three cases (three schools) and then I drew a single-set of cross-case conclusions.

Researcher's Role

I am currently a high school principal in a district with 8800 students overall and 2600 students at my school. I have been the principal at that school for the last 11 years and the assistant principal at the same school for the prior nine years. My 20 years in administration has allowed me the opportunity to form great professional relationships and to share with colleagues my educational thoughts and structural strategies. I did not use my district in this research study, and this allowed me to approach the participants with more of a blank slate than I would have had with my same-district colleagues, given my established relationships there. I needed to and did enable openness and trust by 1) informing all participants of the study, 2) not using deceptive practices, and 3) sharing my research information with the participants (Creswell, 2013).

Site and Participant Selections

I conducted interviews at three different Indiana high schools, Washington High School, Jefferson High School, and Lincoln High School.¹ Each school used a different one of the three evaluation models (RISE, TAP, or PAR). My selection of the three school districts and the specific high schools was based on personal knowledge of the evaluation model used by various schools resulting from my membership in professional education organizations. I considered schools that I knew to be using each of the three evaluation models and determined which schools had been using each particular model for the previous five-year period. I was granted

¹ Pseudonyms are used throughout to represent the three schools at which I conducted the interviews.

PERCEPTIONS OF TEACHER EVALUATION MODELS

permission to conduct interviews at three schools from Ball State University's IRB. Each of those schools agreed to participate in my study.

Each of the three high schools studied serves students grades 9-12. Washington High School is an urban school with 522 students and 25 teachers. Jefferson High School is a suburban, township school with 2,298 students and 105 teachers. Lincoln High School is a school in a rural community with 187 students and 15 teachers. The sample consisted of enough stakeholders to experience prime learning about the phenomenon through balance and variety of participants (Stake, 2005).

Table 2

Comparison of the Three High Schools

	Number of students	Teacher/Student Number and Ratio	Demographics
Washington High School	522 students	35 teachers: 20:1	Urban high school
Jefferson High School	2,298 students	105 teachers: 22.5:1	Township school located outside a large city
Lincoln High School	187 students	15 teachers: 21.4:1	Rural community high school

I interviewed an administrator and two teachers at each high school for a total of nine participants.² According to Creswell (2013), five to 25 individuals who have experienced the subject matter beings studied is an appropriate number for a researcher to include in the study, so my nine participants was well within that appropriate range. My goal in selecting participants

² Pseudonyms are used for each of the nine participants. To facilitate remembering which names were associated with which schools when all nine participants are discussed together in Chapter 7, names beginning with A, B, and C were assigned to the first school, names beginning with D, E, and F were assigned to the second school, and names beginning with G, H, and I were assigned to the third school. Also, alternating pronouns were used to further protect anonymity.

PERCEPTIONS OF TEACHER EVALUATION MODELS

was to identify those who would be “key informants,” who were able to provide a breadth of knowledge rather than looking to find similar or identical representative viewpoints. Also, as Mears (2009) pointed out, a random sample is not the objective in qualitative research but rather the goal is to “interview a small number of people who know a great deal about the topic” (p. 89).

During my initial conversations seeking approval for my study with each of the three districts’ superintendents, I was also able to discuss and choose the administrators I would be interviewing. I wished to interview the administrator that had the most experience using the evaluation model at each high school. In each conversation, we agreed on that person, and I was given permission to contact him or her. I contacted each administrator to introduce myself, explain the nature of my contact, and seek his or her agreement to participate in my study, and all three administrators agreed to participate. The Washington High School administrator, Carter, had four years of experience with the RISE model. The Jefferson High School administrator, Freda, had five years of experience with the TAP model. The Lincoln High School administrator, Ian, had four years of experience with the PAR model.

In my initial contact with each administrator, I also asked about which teachers would best fit my selection criteria. I wanted to interview at each school a new teacher, defined as one having five or fewer years of experience at the school, and a veteran teacher, defined as one having six to twenty years of experience at the school. I sought and was given permission to contact the teachers at each school, and all six teachers agreed to participate in my study. At Washington High School, the new teacher, Adam, had five years teaching experience and the veteran teacher, Betty, had 12 years teaching experience. Both had been evaluated under RISE for five years. At Jefferson High School, the new teacher, Danielle, had five years teaching

PERCEPTIONS OF TEACHER EVALUATION MODELS

experience and the veteran teacher, Evan, had 17 years teaching experience. Both had been evaluated under TAP for five years. At Lincoln High School, the new teacher, Gary, had four years teaching experience and the veteran teacher, Hillary, had 16 years teaching experience. Both had been evaluated under PAR for four years. This arrangement allowed for three viewpoints of each of the teacher evaluation models: one administrator, one new teacher, and one veteran teacher.

Table 3

Comparison of Administrator and Teachers

	Teachers		Administrator
	Years' experience teaching		Years' experience with model
	<i>New teacher</i>	<i>Veteran teacher</i>	<i>Administrator</i>
Washington HS	Adam 5	Betty 12	Carter 4
Jefferson HS	Danielle 5	Evan 17	Freda 5
Lincoln HS	Gary 4	Hillary 16	Ian 4

Data Collection Techniques

The interview protocol I designed (Appendix A) was the instrument for collecting the data on teacher and administrator perceptions of teacher evaluations (Creswell, 2013). These questions about the teacher evaluations were formed from prior experiences, knowledge gained through university teacher preparation programs, and the research conducted as a part of this study. After determining my overarching research question, I developed three more detailed research questions that were intended to cover the sub-categories of information needed to answer the overarching question. Then, I designed the individual interview questions to fit within those three categories. In the end, I used six interview questions to gather data to answer the first research question pertaining to professional growth, four interview questions to gather

PERCEPTIONS OF TEACHER EVALUATION MODELS

data to answer the second research question pertaining to student achievement, and four interview questions to gather data to answer the third research question pertaining to teacher and administrator perceptions.

In August 2017, I scheduled interviews to be held in September and October 2017 with the administrators and teachers at the participating schools, and I emailed each participant the interview questions (Appendix A) one to two weeks prior to his or her interview. I also sent them all consent and voluntary participation forms.

I conducted 60-minute interviews with the administrators and teachers individually in each of the respective school settings. I met with the participants at their schools, in the teachers' classrooms or administrator's office. I scheduled the interviews with the administrators at the time that was most convenient for each individual. Generally, I scheduled the teacher interviews before or after school; however, I also used teacher preparation periods. In keeping with a qualitative study, I pursued a personal connection based on shared educational experiences with the participants in order to mine data that was rich in its anecdotal value.

To start each interview, I collected the completed consent and voluntary participation forms and reminded the participant that I was going to record the interview. I reviewed anything that the administrator or teacher had with him or her (e.g., district policies, checklists, personal records) and established a comfort level between us with preliminary, small talk. According to Merriam (2009), qualitative research uses personal interviews because it "requires a data collection instrument that is sensitive to underlying meaning when gathering and interpreting data," (p. 2). Remembering this, I made sure the interviews allowed me to gather in depth responses, ask follow-up questions in real time, and gain additional body language information. I followed my interview script, asking follow up or clarifying questions only when necessary.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Because the interview was recorded, I did not take notes on the participants' answers; instead, I was able to take very short field notes during the interviews to memorialize contextual information, such as body language, and I added to those field notes after the interviews, intending to include these notes in my data analysis. The participation was voluntary, and the participants were assured of the confidentiality of their information and that it would only be used for the sole purpose of this qualitative study.

At the conclusion of each principals' interview, the principal and I walked around and observed the teachers in would later interview in their classroom setting to help me gain better insight into the classroom practices.

Managing and Recording Data

I kept all of the documents that I sent to each school in preparation for my interviews in folders labeled by school. I also kept all documents (district policies, checklists, personal records) I received from each interview in its own folder. After each interview was completed, I had the interview recording transcribed. I went through the copy of the transcript with the recording and verified the accuracy of each transcription. I kept the transcribed interview in the folder for each interview.

Seidman (2013) suggests keeping the generating and analyzing of data separate; in fact, he recommends avoiding any in-depth analysis of the data until all interviews are completed. This decreases the chances of imposing meaning from one participant's interview onto later interviews. I followed this practice so that I could "minimize imposing on the generative process of the interviews what I think I have learned from other participants" (2013, p. 114). Once I had all nine interviews completed, transcribed, and checked over for transcription accuracy, I began my analysis of the data.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Data Analysis Procedures

While the interviews were being transcribed, I reviewed in detail the documents the participants gave me during their interviews. I paid particular attention to those documents that outlined and described the components of the evaluation models and made notes to be used in discussing my findings. Once the interviews were transcribed, I organized with the transcripts all related data. This included my observational field notes and the evaluation documents. I then reviewed all of the data and looked for concepts or themes. To do this, I read each transcript line-by-line. This process of reviewing data is called open coding, and it required me to open all of the data to all potentials and possibilities contained within them (Corbin & Strauss, 2007; Saldaña, 2009). According to Saldaña (2009), “A code in qualitative inquiry is most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attitude for a portion of language-based or visual data” (p. 3). As a result, the codes allowed me to identify major themes that emerged from the interview and observation data and, after analysis, draw conclusions based on those themes. The data was analyzed using the interpretational analysis method. This was an iterative process of individual and group level (i.e., grouping by schools) review of the information collected. Additionally, relying on Seidman (2013), I recognized the importance of not addressing the data with any pre-conceived notions or hypotheses to which I was trying to match the data. I came to the data “with an open attitude, seeking what emerges as important and of interest from the text” (p. 117).

Coding is not a precise science, it is primarily an interpretive act that allows the researcher the ability to summarize data, not reduce it (Saldaña, 2009). This allowed me to determine any patterns and themes within the teacher and administrator interviews that emerged to explain perceptions of teacher evaluations. I analyzed the transcripts one at a time, looking for

PERCEPTIONS OF TEACHER EVALUATION MODELS

key words and phrases used by the participant to answer my questions. I largely followed the approach outlined by Seidman (2013) whereby I searched each transcript for “connecting threads and patterns” (p. 125) and highlighted those words, phrases, and chunks. Then, I reduced each interview to a list of those highlighted parts looking for common subjects so that I could create categories and themes. This identification of emerging themes allowed for data to be extrapolated in order to provide a basis for comparison between participants at the same school.

Once the themes and categories within the case analysis of the three schools were developed, then analysis of the themes across the cases (cross-case analysis) was done to help determine whether the evaluations affected professional growth and student achievement, and what influences the evaluation process and outcomes. According to Saldaña (2009), this is commonly known as second cycle coding methods, as it is a more patterned or focused coding (p. 48). During this phase of the coding, I color coded the data that I found meaningful so that I could more easily make the necessary comparisons between the three schools and evaluation models. I largely repeated the process I had undertaken with the individual interviews; however, this time, I was comparing the data categorized from the three schools to determine patterns and themes that were common across the three different evaluation models.

Throughout the process, I relied on Seidman’s (2013) teaching that I, as the researcher, must be confident in the role of my judgment in identifying what is important from the interview transcripts. My judgment, in accordance with Seidman, was informed by past experience and by working with and internalizing the interviews themselves and is the most important aspect that I brought to my study. In the end, my synthesis of all of the data and judgment as to commonalities and themes led to me drawing conclusions that ultimately provided the answers to my research questions.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Limitations

As with any research there were limitations to this qualitative study. Marshall and Rossman (2011) defined limitations as a reminder to readers of what a study is and what it is not, and its boundaries and results must be understood in that context. The qualitative research for this dissertation was grounded in the interpretive theoretical perspective (Creswell, 2013; Hoy, 2010; Lichtman, 2013); thus, the greatest limitation of the study was that human interpretation and the use of self-reported information through interviews was the basis for developing the knowledge regarding perceptions about teacher evaluations. Although I assumed that respondents would be honest, there was the possibility that answers would be given to portray the school district or the evaluation process at that school in a better light than they really were.

I felt that another limitation might spring from my overall experience and biases regarding teacher evaluations. I have been a teacher and administrator for over 25 years in three school systems, which have used many different evaluation models and techniques. This could have been a limitation because I have already-formed my perceptions and experiential information as to when teacher evaluations work and when they do not work. Without care, this could have led to bias in my questions and potentially in my evaluation of the data gathered. Thus, I paid particular attention to maintaining my neutrality and was careful to approach both my interactions with participants and my evaluation of data as objectively as possible. This increased difficulty in evaluation of data is, of course, a known limitation to qualitative research generally.

Ultimately, the primary limitation will be that the research questions combined three very complex phenomena: (a) teacher evaluations; (b) the amount of teacher growth in the classroom (from the teacher and evaluator perspectives), and (c) the degree of actual increased student

PERCEPTIONS OF TEACHER EVALUATION MODELS

achievement, and I was not able to comment on all of the research nor perceptions that speak to these complexities. As stated by Thorne (2016), “[w]hat remains, then, are the thoughts of one individual scholar, presented in the context of a smattering of the relevant background to position these thoughts within a larger context of evolving ideas, and a good deal of experiential perceptions” (p.15).

Summary

The qualitative research that I conducted focused on naturalistic inquiry by means of my getting the necessary permissions and then personally interviewing (a) an administrator and two teachers from Washington High School that uses the RISE evaluation method, (b) an administrator and two teachers from Jefferson High School that uses the TAP evaluation method, and (c) an administrator and two teachers from Lincoln High School that uses the PAR evaluation method. This methodology allowed me to use the natural setting to gather information without my control or manipulation. Data was in the form of the transcribed interviews and my personal observations. I analyzed the data individually and comprehensively and used open coding to identify themes and draw conclusions from those themes. The following chapters (4-8) will be done in case study form with an analysis of three case studies. These multiple case studies allowed me to analyze across cases in order to answer the research questions.

PERCEPTIONS OF TEACHER EVALUATION MODELS

CHAPTER FOUR: FINDINGS – CASE STUDY #1

The purpose of this qualitative case study is to analyze three High-Stakes Accountability Models--RISE Evaluation and Development Systems (RISE), Teacher Advancement Program (TAP), and Peer Assistance and Review (PAR), and this chapter reports on a Washington High School, which is part of a school district using the RISE evaluation system. As previously explained, data was collected on teacher and administrator perceptions concerning: (a) the general methodology or structure of the three evaluation models; (b) the specific practices for gathering and compiling the raw data in each of the three models (i.e., the forms used); and (c) the potential impact of each model on teacher effectiveness, student achievement, and professional growth.

At Washington High School, I interviewed Carter, an administrator with the most knowledge about and experience with the evaluation model used by the school; Adam, a teacher with five or fewer years of teaching experience; and Betty, a teacher with six to twenty years of teaching experience. This chapter will describe both the school profile and the evaluation model used at the school, background on the participants, and the themes that emerged from the interviews.

Washington High School: Profile

WHS serves 522 students in grades 9 through 12. It is located in an urban community situated close to a large city. The demographic breakdown of WHS is shown in the table below.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Table 4

Demographic Breakdown of WHS

WHS student category	Percentage
Special education	13
Free and reduced lunch	51.9
English language learners	5.6
White	56
Black	21.6
Hispanic	11.7
Multi-racial	5.6
Asian	5.1

Carter, the WHS administrator, described the following about WHS during his interview:

Traditionally, 80%-90% of WHS students matriculate to post-secondary education, including vocational schools, two-year associate degree programs, and four-year bachelor programs. A majority of WHS students take the ACT as a college entrance exam, an increasing number of students take Advanced Placement exams, and the graduation rate has been over 95% for several years.

The curriculum addresses a wide range of post-secondary goals, and the school offers vocational programs. WHS traditionally graduates more than 80% of the students with a Core 40 and/or Academic Honors Diploma. The school offers 11 Advanced Placement courses and two College Credit courses. WHS makes a comprehensive curriculum guide available online, and the guide illustrates what each course has to offer along with the basic requirements for graduation, the various types of diplomas, and a four-year planning guide. In addition, basic enrollment requirements at WHS require all students to enroll in an English, math, science, and social studies course each semester they attend. The vision for WHS is to prepare students for college and workforce readiness in the 21st Century. Its mission is to inspire and prepare students to pursue excellence academically, emotionally, and socially. The school core values are as follows:

PERCEPTIONS OF TEACHER EVALUATION MODELS

1. Academic

- Rigorous curriculum aligned with Indiana Academic Standards and Common Core Standards
- Engaging classroom instruction based on researched “Best Practices.”
- Authentic assessments which measure student progress and achievement
- Integration of technology in the classroom

2. Emotional

- Provide a safe and comfortable environment
- Relationship building
- Management of conflict resolution

3. Social

- Personal Integrity
- Importance of attendance and timeliness
- Communication skills
- Cultural awareness and appreciation

Washington High School: RISE Evaluation Model

According to Carter and the school evaluation plan documentation he provided me, WHS adopted and has been using the RISE Evaluation and Development System for evaluating certified staff since 2012-2013 school year. Prior to that, all that was in place was a goal setting system for evaluating teachers. According to the district handbook, the RISE teacher evaluation process was chosen to recognize great teaching and to support growth in each teacher. The district states that every teacher, regardless of experience deserves meaningful feedback on their performance on an annual basis. It believes that RISE is student focused and uses multiple

PERCEPTIONS OF TEACHER EVALUATION MODELS

measures to show teacher performance over time. WHS teachers are evaluated by administrators who have completed training and given support in evaluation. The administrators are responsible for evaluating, tracking staff evaluation results, approving Student Learning Objectives (SLOs), and helping the staff to set goals for professional development.

At WHS, an administrator has an orientation session with a new staff member to explain the process. Each new teacher is enrolled in the Pivot software suite, a tool to (a) collect, code, rate, and reflect on evidence, (b) communicate securely, (c) set goals for improvement, (d) track professional development, and (e) monitor progress toward student learning objectives. The weights for the various parts of the RISE Evaluation used at WHS are 75% for the classroom observation using the Rubric evaluation tool, 20% for SLOs, and 5% for school-wide letter grade. Each teacher receives a rating at the end of the school year in one of the four performance levels: Highly Effective, Effective, Improvement Necessary, and Ineffective. Assessment of instructional knowledge and skills that influence student learning are measured by the competencies in the RISE Teacher Effectiveness Rubric. All teachers are evaluated in the domains of Planning, Instruction, Leadership, and Core Professionalism. The evaluation system for the year includes three short (10 minute) evaluations and two extended (entire class period) evaluations. The three short evaluations are random pop-ins, at any time in the school year. The two extended observations are scheduled and done on days and periods chosen by the teacher. Both short and extended observations are a time for administrators to collect information. There is no summative rating assigned until all information is collected and analyzed at the end of the school year.

All WHS teachers write SLOs. Assessments for measuring the success of meeting the SLOs are centered on ISTEP, ECA, and AP testing results. In subject areas where state

PERCEPTIONS OF TEACHER EVALUATION MODELS

assessment or district standardized tests do not exist, teachers work with administrators for approval of assessments and targets. When WHS receives its district rating, the data is entered into PIVOT and will contribute to each teacher's summative evaluation. Feedback is given via electronic transmission of observation results and a conference between the administrator and teacher. There is also a summative post conference after the fifth observation.

WHS uses the consequences described in the RISE model for any teacher who negatively affects student achievement and growth and does not receive a rating of highly effective or effective. When this should happen, the administrator and the teacher develop a remediation plan of not more than 90 days in length and correct the deficiencies noted in the teacher's evaluation. Finally, a teacher who continues to receive negative performance evaluations can request a private conference with the superintendent.

Washington High School: Participants

Carter, who has been at WHS for four years, had been at other schools prior to coming to WHS but has only evaluated teachers using the RISE model. Prior to using RISE, he evaluated teachers without a specific model. Adam's five years of teaching experience have all been at WHS. Adam has taught in the same department his entire tenure and has taught various courses from introductory to Advanced Placement. Betty's 12 years of teaching experience have also been at WHS. She has also taught in the same department, teaching introductory to Advanced Placement courses, her entire career; however, she has taught at the same grade level for the last several years.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Washington High School: Themes that Emerged from Research Question 1

The first set of themes emerged as a result of data collected from participants' responses to the first research question: What characteristics of the RISE teacher evaluation model do teachers and administrators perceive to be the most effective in promoting professional growth? Four themes emerged: RISE caused changes to instruction; RISE generated valuable feedback that was incorporated into instruction; RISE included beneficial training; and the timing of certain RISE components necessarily affected the timing of growth.

Theme 1: RISE caused changes to instruction

The first theme that emerged was that RISE caused teachers and administrators to change instructional practices. Carter believed that RISE caused him to change his observation methodology. The rubric helped him pinpoint good teaching techniques and pushed him to be a more focused and purposeful instructional leader. Carter said, "I believe that each year has helped me get to the idea that I could pinpoint instructional strategies quicker and give feedback in a more effective manner and a more efficient expected time frame." Adam also felt that the RISE model had changed his teaching methods. He said, "Using the RISE system, I moved from a middle-of-the-road teacher who could control a classroom of students to a teacher who felt confidence in my abilities to challenge students and reach my students at a far different level" and that "[b]ecause of the system, I knew what I was looking for as a teacher and I knew what I needed to do to get better." Adam said that specific competencies caused him to change the way he planned and even the way he phrased questions in class. All of this caused Adam to see professional growth, especially when he used the RISE rubric to identify areas to grow and then to track exactly where growth had, in fact, occurred.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Similarly, Betty, who thought she was a good teacher before RISE, still admitted that the model has given her good ideas to implement, and that it “has helped [her] not get lax and stagnant with [her] teaching.” Betty cited an example of reincorporating a good practice (cold-calling on students) when she was being observed because the RISE rubric helped her to recognize that she had let doing so slide a bit. Betty also mentioned changing her behavior so that student learning was more evident in observations and so that she was using data more effectively. She said, “I have been introduced through the RISE system to using numbers to drive how I do what I do.”

Relatedly, across the board, the participants agreed that their preparation for class has improved under RISE. Carter admitted to being “a little old school,” as he emphasized the importance of a great lesson plan to an effective learning environment; Carter referenced domain 1 in the RISE model, which “spells out how to use the data to plan out lessons and have the end in mind on what the students will achieve and what was learned when the class is concluded.” Carter added his belief that for the specific competencies in the effective learning domain to happen in class, teachers must plan the details in their lessons. Adam talked often of this sort of preparation and even went a little further when he said, “so for every lesson I always plan extra activities in case we get through something quicker than expected or if we need a different change of pace” because something planned did not work as it was intended to work. Similarly, Betty discussed planning at least three to four separate activities for any given class as well as the transitions to move between those activities. This was specifically done to be more purposeful about engaging the highest number of students, which all saw as evidence of better teaching.

PERCEPTIONS OF TEACHER EVALUATION MODELS

The common language that was consistently used in the RISE model also made it easier for teachers to improve their instructional strategies. All three participants believed that consistently using the PIVOT software (a data management and evaluation “warehouse” tool) was very helpful. They all reiterated the fact that there have not been any changes nor alterations to the evaluation process over the past five years. The interview responses showed the RISE model allowed for positive growth because evaluators and evaluatees were using the same set of components and terms to describe the identified teaching attributes. As Adam said, “The administrator and the teacher are on the same page,” and both understood “what was expected to be seen and shown in both the short and the long evaluations because the model provides a common language and vocabulary that reflects the realities of a classroom.” He said that “[t]he rubric pushed me to get better in the categories because I was seeing the positive effects in my class.” Betty added, “The RISE model offers consistency across all disciplines. Consistency of administrators using the same evaluative tool and consistency in the administrator who evaluates you from year to year have all helped to promote professional growth.” Carter specifically mentioned pointing teachers to the rubric language and celebrating with them when through the course of a year or years they saw growth as they kept working to implement consistent components.

Theme 2: RISE generated valuable feedback

The second theme related to promotion of professional growth was that the feedback component of RISE was very valuable. Adam thought that one of the most important parts of the RISE model were the administrator pre- and post-conference meetings. Under RISE pre-conferences were typically done and post-conferences were required so scheduled between administrator and teacher after every observation. It was clear that all three participants believed

PERCEPTIONS OF TEACHER EVALUATION MODELS

that taking that time to sit down and have professional conversations contributed to professional growth. Adam said that when the post-conference meetings were held he knew that “coaching was going to be the main theme. The observation would be critiqued and then open-ended discovery questions would be asked to enhance dialogue.” Betty recalled being given things to think about in post conferences that she later implemented to benefit student learning. Carter indicated that post conferences, which he described as “the most critical part” allow each participant the opportunity for collaborative, reflective conversations with each other that lead to a heightened awareness of other people’s perspectives. Carter noted that “[t]he feedback and dialogue between the administrator and teacher provide a common vision and expectations for the classroom and doing the job well.” He even said that the RISE expectations for these conferences was the thing that helped him grow the most.

Carter and Adam both felt that the RISE evaluation model allows teachers and administrators to work well together, and because of that working relationship, professional growth was developed through trust and openness in the collaborative process. Adam said, “Trust has been set up with our evaluations when the administrator observes you; they are primarily giving you positive feedback or critiquing your lesson and saying things like ‘I like what you have done here, or I do not like what you done here.’” Adam described a post-observation discussion where he was offered a critique, but given the chance to interact with the evaluator. He concluded that this practice contributed to what he saw as growing “leaps and bounds as a teacher.”

Carter was a believer in this growth mindset, and he saw the multiple evaluations performed with RISE as both contributing to his growth (“by watching other great educators in action”) and allowing him to help others grow. Specifically, he said, “I have become a much

PERCEPTIONS OF TEACHER EVALUATION MODELS

better evaluator because my arsenal is larger and wider because I have seen other things and can predict where problems and growth areas will be needed.”

Correspondingly, all three participants believed that WHS very purposefully did not employ the RISE model in a punitive manner and this purely positive manner of feedback contributed to growth. Adam, Betty, and Carter were all very clear on this point. They all mentioned this more than once, related to different sub-questions. Adam explained how WHS evaluations were not set up “to be a ‘gotcha’ system” and explained how the RISE system as used at WHS allowed a teacher to give something a try, have it be observed, and get feedback on it. Betty made it clear that administrators dealt carefully with teachers (especially long-time veterans) by alleviating the fear and angst that might come along with having someone in your classroom, and she characterized RISE as a growth, not punitive system, “a way to get better.” Carter mirrored and affirmed what his teachers said. He even talked in terms of negative or “constructive feedback” coming in terms of “what we can do together to make it better.”

Theme 3: RISE included beneficial training

The third theme that emerged was that RISE involved training that was beneficial. Full staff training on RISE occurred during its first year of implementation, according to Carter. In fact, the full staff went through and trained on each of the indicators for an entire semester. After that, as Adam explained, first-year (or new-to-the-district) teachers are given a full day of training before the school year began, and all teachers are given additional training during the weekly PLC (Professional Learning Communities) meetings to start each year. Carter confirmed that “each year the training stays fairly consistent as has the format [with] much of the training occur[ring] early in the year through the work in PLCs to refresh and get everyone back on the same page.” He also added that the training involves both understanding the indicators

PERCEPTIONS OF TEACHER EVALUATION MODELS

themselves and understanding the research behind the model; in fact, “*Teach Like a Champion* and other research books have been brought in to enhance the understanding of the model.”

Betty recalled that the training her first year was especially good because “the administrators themselves had gone to outside training, bringing back that information to the staff, to say here’s why we’re doing what we’re doing and how we are going to do it here.” The teachers then worked in small groups on how to do SLOs and TLOs (targeted learning objectives) and what theirs should be. Adam said that because the rubric has stayed consistent, he felt that once a teacher was trained, it was fairly easy to log-on the system and see the expectations and how those expectations are being met.

Betty said that in addition to the formal training, WHS also had a system of informal training. She mentioned that department heads routinely trained newer teachers on an as-needed basis and that other teachers generally “just help[ed] out peers.” In fact, she talked about a time with a new teacher where they “had a common prep, so he was able to come ask me questions like how to [write SLOs].” Betty questioned whether this “would happen at a school bigger than ours” but concluded that she “would like to think helping a colleague is helping colleague no matter where.”

Adam felt that there was the “just enough” training and that teachers who were acting professionally learned the system and used it to improve. Similarly, Carter thought the training was appropriate and helpful. He concluded that training “each year does allow us to discuss focus points for the year with evaluations based on our school data or any changed requirements from the state or federal level. This helps to inform the whole teaching staff on where we’re strong on our school data and where our areas of growth lie and how this related to the [RISE] indicators.”

PERCEPTIONS OF TEACHER EVALUATION MODELS

Theme 4: Timing of certain RISE components affected timing of growth

The fourth theme that arose regarded the timing of the evaluation system as being a component that affected when teachers could implement growth strategies. This came from two related aspects: certain SLOs and the school letter grade, which are both components of the summative final evaluation. Recall that the SLOs were worth 20% of the total and 5% of the total was allocated to the school letter grade. For all teachers, the school letter grades are not released until the fall of the next school year. As Carter said when talking about this, RISE

is designed to allow teachers to get specific feedback to improve, but the whole part of waiting for the information to finally get there to finish up their summative grade for the year was just disappointing to many on staff.

Also, both Adam and Betty taught AP classes and had SLOs tied to AP test scores, which are not released until July, so both end up starting the new school year with only 75% of their previous year's evaluation known to them. Betty talked about working on her SLOs for a new year but not yet having had an individual conversation about how she did on her SLOs from the previous year; she saw this as a "bad thing." All participants saw this a problem with the model, but none was able to suggest a solution.

Washington High School: Themes that Emerged from Research Question 2

The second set of themes emerged as a result of data collected from participants' responses to the second research question: What characteristics of the RISE teacher evaluation model do teachers and administrators perceive to be the most effective in promoting student achievement? Two themes emerged: teaching became increasingly student focused, and data was being used to benefit students.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Carter summed it up well when he said the RISE model allowed for discussions with all teachers that “made them believe that great student achievement is possible and productive, and because of this, we get good results and have high expectations of each other.”

Theme 1: Teaching became more student-focused

The first theme that emerged was that the evaluation model led WHS as a school to become much more concerned with what students are doing and how they are learning and less on what was easier for the educators. Under RISE, a central competency for instructional strategies focuses on student engagement and checking for understanding. Both teachers indicated that they were much more student-centered than pre-RISE, and both have changed from mostly lecturing, sit-and-get, to a more discussion style, hands-on learning atmosphere. Adam stated his predominant instructional strategy is to know the kids ,and, while acknowledging that the material is the same, relying on the RISE rubric has caused him to change how he is teaching based on who he is teaching. He said, “I use [knowledge of my students] to create my activities, I use it to create my checks for understanding, I use it to form my formative and informative assessments.”

Betty talked about how RISE had forced her to more specifically consider student learning in the strategies she focused on, giving an example of time spent (pre-RISE) on a class website that contributed little to student achievement and how that time now went to delineated competencies. Carter made it clear that he used the effective learning competencies to pinpoint exactly what teachers are and are not doing and that he “can clearly see which teachers are . . . highly effective [with regard to student engagement and checking for understanding] and where specific RISE model indicators can help facilitate growth.”

PERCEPTIONS OF TEACHER EVALUATION MODELS

Theme 2: Data used to benefit students

The second theme that came out the interviews was that RISE requires use of student data and that when discussions about using data occur they are more productive and increased student achievement is more likely. All three participants stated that the model requires data to be used in the evaluation process, and all three believe they are better for it doing so. In fact, for Betty, this is the primary area that she sees RISE as making a difference in her teaching, which she sees as translating to making a difference in student achievement. Teachers are required to have SLOs, and administrators expect them to track their data and report it. Betty stated that she did not previously (pre-RISE) use data in any meaningful way but that because she “had to” under RISE she did and once she saw the benefit, she naturally incorporated data into her teaching, seeing the results in her classroom. In fact, the RISE model pushes her to have goals and expectations that are clear and measurable, and she feels it allows administrators to work with teachers so that their goals are appropriate and attainable while providing a realistic expectation of improvement.

Carter alluded to the fact that “the data to drive instruction has greatly improved under the required SLO portion of RISE model. This has been huge because the staff continues to improve on their data and their ability to track data.” Adam commented,

When I went through college, we did not talk about data at all. That was not something that was discussed. But now, I use data for everything to help drive my instruction. . . . Data does not have to be the hard grades you put in the gradebook, data can be from formative assessments. Looking around the class, you can see that 75% of the students got the answer right. I might then change or work with some groups of students a bit differently to reteach to the students who were not sure of the material.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Adam also attributed having “grown in [his] area of expertise” to the RISE model requiring the use of SLOs and data. Given the data portion requirement, WHS has aimed to ensure that teachers perform at their best to enhance student learning. This led to a greater sense of accountability under the RISE model.

Washington High School: Themes that Emerged from Research Question 3

The third set of themes emerged as a result of data collected from participants’ responses to the third research question: In what ways do teacher and administrator perceptions of the RISE evaluation model influence the evaluation process and educational outcomes?

Two themes emerged: teachers who appreciate getting feedback and younger teachers look favorably on the RISE system, and there were two negative perceptions concerning RISE, but they were not serious enough to influence the process or outcome.

Theme 1: Most teachers looked favorably on RISE

The first theme that emerged is that teachers who view getting feedback as a positive thing and those who are younger (or who had only ever been evaluated under RISE) are likely to view RISE as a positive thing. This was first illustrated by both Adam and Betty saying they appreciated getting feedback and feeling that it was a positive part of the evaluation process that directly helped them grow and affect student achievement. As for Carter, the conversations that occurred as a result of the RISE feedback requirement were also perceived incredibly favorably (his favorite aspect, in fact). Carter confirmed that “[WHS] always got back that [teachers] enjoyed the feedback aspect of the model and a clear message of expectations for evaluations.” Thus, to some degree it seemed that all three participants’ perceptions of the benefits of feedback contributing to influencing them to be more positive about RISE than other teachers less interested in feedback might be.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Adam admitted that he did not need someone watching over him all of the time, but said “there are so many little things you can do better if you had someone watching you more or conversing with you about it more.” It was obvious that his belief in this regard contributed to his positive view of RISE. Betty added, “I realize constant reflection practice is an expectation of the RISE model. I, myself, am taking that feedback as a positive consequence of being evaluated and internalizing it as a part of my bigger push to get better.”

Also, it seemed that all three participants thought that the newer teachers, or perhaps those who had only been evaluated in under a RISE model, were more likely to respond positively to the evaluation system than more senior teachers (those who had taught for years without an evaluator in their classroom five times throughout the year). Betty said that she sometimes “just wonder[ed] though are other teachers just listening to the advice and only doing it because they know they are being evaluated; they don’t really buy in.” She also thought that “many teachers feel it is more of another thing to accomplish on the yearly task list” and give it little additional thought. Carter acknowledged the differing perceptions of RISE, and Adam spoke to this specifically to this phenomenon. He added,

[t]he older, experienced teachers just see it as a game . . . [and] do not respect the evaluation system, saying “I have taught for a long time and I do not need this.” Or “You’re are not going to tell me anything I do not already know.” But for a newer teacher who wants feedback, wants to grow, and who wanted to improve, the RISE system is excellent.

Theme 2: Negative perceptions did not influence overall processes or outcomes

The second theme centered on two negative perceptions about RISE and the fact that the presence of these negatives did not cause any of the participants to doubt the efficacy of the

PERCEPTIONS OF TEACHER EVALUATION MODELS

RISE model. First, the participants expressed concern that an evaluator could have a difficult time gaining a true understanding of a teacher's effectiveness in the limited time spent in the classroom. All acknowledged that six visits on the one hand seems like a lot, yet that it is only observing approximately two-and-one-half hours of instructional time during an entire year. As Adam put it, "How do you even go about encompassing a years' worth of teaching by two-and-a-half hours of observation?" He felt that a teacher could do anything well for two classes but that "getting 4.0's on those two observations and then being a lazy teacher for the rest of the year did not seem right or equitable." Betty echoed this concern. In fact, it caused her to say that the unannounced visits had more validity because "[a]nyone can put on a great show for 50 [scheduled] minutes." She went so far to say that even the unannounced visits were questionable for predicting with certainty a teacher's day-in-and-day-out ability. Carter was not as obvious, but he did make it clear that he "would not want to reduce the number of evaluations" and that in general the more chances an evaluator had to be present in a teacher's classroom the better.

Closely related to this was both Adam and Betty wanting feedback from peers and finding fault with the evaluation model for relying only on the administrator observations. Adam wished "there was more time and expectations to watch other teachers teach and do more collaboration on a professional level than the RISE model promotes." Betty also commented,

I think a challenge is that there is no peer component to the RISE model. Maybe a non-evaluative piece but just more peer feedback, because you just never go in to other teacher's classrooms, and I think it would be interesting to get peer feedback. It is one thing to have an administrator come in that has maybe been out of the classroom for ten years versus someone who was just teaching coming in to give me feedback.

PERCEPTIONS OF TEACHER EVALUATION MODELS

While Carter did not speak to this directly in terms of a perceived problem with RISE, he did speak often about the benefits of feedback from multiple sources and specifically mentioned how much he has professionally grown by “getting to watch other educators in action.”

The second negative aspect that all three participants addressed had to do with potential biases. At a base level, Carter mentioned his belief that multiple evaluators could provide “inconsistencies in narratives and conferencing based on the administrator’s skill set,” meaning that different evaluatees could hear different things from different evaluators. He also talked about how an evaluator’s personal circumstances (what else had happened that day or week) and personal biases might “get in the way” of an unbiased evaluation.

Adam and Betty both used the exact wording of “preconceived notions” when talking about potential evaluator biases. Betty said she thought that evaluators typically held a belief that a teacher was or was not good and they went into an evaluation expecting to see something that aligned with what they already believed. Adam agreed and added that this belief may have incorrectly been formed by “kid input and sort of the good ole boy network” even though neither may be an indication of best practices teaching. Betty also felt that the reality that WHS is a high performing school with a self-professed expectation that it hired only “highly effective teachers” often meant that teachers were labeled as highly effective because “they have to be to be working here.” This might lead to a bias that evaluated merely effective teachers more highly than they should be. Even with these “negatives” noted, it was still apparent that all three participants believed the process used by RISE and the outcomes garnered from RISE to be beneficial for teachers and students alike.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Washington High School: Summary

In this chapter, the findings from Washington High School (Case Study #1) were reported in order to answer the specific aspect of the overarching research question relevant to WHS: from the viewpoints of both teachers and administrators, does the RISE teacher evaluation model actually, in practice, cause teacher growth and improve classroom instruction? The findings of the three interviews were coded and themes were determined as they applied to the study's three research questions. With regard to research question #1 pertaining to characteristics of the RISE model that the participants identified as most effective in promoting professional growth, four themes emerged. First, all participants felt that although they were good at what they did prior to RISE, the model made them better. This was seen both in how they prepared and in how performed. Additionally, the common language that RISE brought to the equation was lauded as being valuable and part of what made the model contribute to growth. Second, RISE was seen as generating valuable feedback and that feedback was uniformly incorporated back into instruction, and this directly affected growth. Third, the participants thought the training aspect of RISE was beneficial. Initial training, follow-up training, and collaboration between staff as informal training were all seen as contributors to growth. Lastly, there was a perception that the timing of certain RISE components, most specifically school letter grades and SLOs with student data coming in after the school year, made it difficult for the snapshot of an individual's growth to happen directly at the conclusion of a given year.

With regard to research question #2 in which the participants were asked to identify the characteristics of the RISE model that were the most effective in promoting student achievement, two themes emerged. First, the participants made it clear that teaching became increasingly student focused under RISE, and that changes were made taking into account specific reference

PERCEPTIONS OF TEACHER EVALUATION MODELS

to and plans for student achievement. Second, at WHS under RISE, data was being used very deliberately to benefit students. RISE is a very data driven model, and the participants felt that this not only led to student achievement but allowed it to be monitored and ensured.

Last, with regard to research question #3, which explored the ways that participant perceptions of the RISE model influenced the evaluation process and educational outcomes, two themes emerged. First, younger teachers and those who appreciate getting feedback look favorably on the RISE model, and because such teachers were in the overwhelming majority, RISE was seen as an overall positive, buy-in had happened, and this caused growth and achievement to result. Finally, there were a couple negative perceptions concerning RISE pertaining to the desire for more peer feedback and the potential for biases to come into play in evaluation systems generally, but the participants agreed that those perceptions did not affect the overall value of or beneficial use of RISE.

PERCEPTIONS OF TEACHER EVALUATION MODELS

CHAPTER FIVE: FINDINGS – CASE STUDY #2

The purpose of this qualitative case study is to analyze three High-Stakes Accountability Models—RISE Evaluation and Development Systems (RISE), Teacher Advancement Program (TAP), and Peer Assistance and Review (PAR), and this chapter reports on Jefferson High School, which is part of a school district using the TAP evaluation system. As previously explained, data was collected on teacher and administrator perceptions concerning: (a) the general methodology or structure of the three evaluation models; (b) the specific practices for gathering and compiling the raw data in each of the three models (i.e., the forms used); and (c) the potential impact of each model on teacher effectiveness, student achievement, and professional growth.

At Jefferson High School, I interviewed Freda, an administrator with the most knowledge about and experience with the evaluation model used by the school; Danielle, a teacher with five or fewer years of teaching experience; and Evan, a teacher with six to twenty years of teaching experience. This chapter will describe both the school profile and the evaluation model used at the school, background on the participants, and the themes that emerged from the interviews.

Jefferson High School: Profile

JHS serves 2298 students in grades 9 through 12. It is a township school located outside a major metropolitan city. The demographic breakdown of JHS is shown in the table below.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Table 5

Demographic Breakdown of JHS

JHS student category	Percentage
Special education	9
Free and reduced lunch	51.7
English language learners	18
White	62.9
Asian	15.1
Hispanic	8.9
Black	7.8
Multi-racial	5.1
Indian	.3

During my interview with the JHS administrator, Freda, she described the following about JHS: While minority ethnic groups total 37.1% of the JHS family, the Asian (15.1%) and Hispanic (8.9%) groups are the two groups that have grown substantially over the past few years. The Asian population's dramatic increase is due to an influx of Burmese (Chin) refugees coming to the school district. These students have provided many challenges in dealing with a large ELL population with very limited English skills.

Not only has the face of JHS changed dramatically over the past ten years, the school has also had its free/reduced lunch population increase over 30% in the last 10 years. Freda stated that JHS welcomes this diversity within its building and is committed to addressing the needs of all students.

JHS has had a graduation rate over 90% the last five years. The increase is attributable to several factors. JHS has started looking at students' schedules more creatively trying to get special education students to receive a Core 40 or Honors Diploma. The school uses a PLATO (Platform for Learning and Teaching Online) Lab that focuses entirely on helping students with a full schedule pursue credit recovery programs and also offers struggling students two alternative education programs: the Jefferson Pathways to Excellence program housed at an off-campus site

PERCEPTIONS OF TEACHER EVALUATION MODELS

and the Fast-track program located on site. Both of these programs are designed to keep struggling students in school.

JHS students are more consistently earning a Core 40 Diploma or Academic Honors Diploma, while the number of General Diplomas are on a decline. The school has phased out courses offered that do not meet the Core 40 diploma requirements. It has increased the number of students taking AP courses and has launched an Early College program. JHS is in the second year of an initiative funded by the Gala Foundation to attempt to steadily increase the number of students taking AP courses, as well as improve the performance of students on AP exams. The funding provided by Gala has enabled the school to train teachers to teach AP courses and provides necessary resources. The culture and set of expectations for students taking AP classes has raised the level of academic rigor generally at JHS.

Over the last six years that the state has assigned grades to schools, JHS has gone from an F Rating to a B rating. The last four years JHS's grade has been a B, but Freda emphasized that within that general category the numerical grade has shown an upward trend (from 3.10 to 3.30).

Jefferson High School: TAP Evaluation Model

During my initial meeting with Freda, I was given a copy of the TAP Evaluation model utilized by JHS. JHS adopted and has been using the TAP Evaluation Tool for evaluating its staff since 2012-2013. Administrators and Master Teachers evaluate any teacher who teaches in an instructional setting (general education classroom, special education classroom, physical activity classroom, laboratory, or performance classroom) regardless of whether the instruction takes place in a large group or a small group setting. The TAP Evaluation Tool consists of evaluation in three (3) domain areas – Instruction, Designing and Planning Instruction, and the Learning Environment. Each domain area is sub-divided into indicators—twelve (12) for the

PERCEPTIONS OF TEACHER EVALUATION MODELS

Instruction domain, three (3) for the Designing and Planning Instruction domain, and four (4) for the Learning Environment domain.

Any teacher new to JHS will receive professional development in each of the 12 indicators in the Instructional Domain prior to any formal announced or unannounced observations during the school year. Observations for returning staff may start after the first two weeks of school, as determined by the JHS's administrative team. The administration will provide a plan of additional support, as needed, for new teachers to learn the TAP Rubric. Special consideration may be given to new teachers unfamiliar with the content areas and any unusual circumstances.

JHS follows the Indiana Code requirement that each certificated employee receives an annual designation in one of the following rating categories—Highly Effective, Effective, Improvement Necessary, or Ineffective. The evaluators use the TAP Evaluation Tool to implement the observation, evaluation, and appropriate rating designation of non-administrative certificated personnel. Also in compliance with Indiana Code, a certificated employee who receives an annual rating of Ineffective or Improvement Necessary is placed onto a remediation plan. JHS believes that evaluating teaching effectiveness is basic to improving instruction and in facilitating professional growth. The performance indicators listed in each teaching domain guide teachers and administrators in the evaluation process and allow administrators to make decisions regarding teacher effectiveness ratings.

A Highly Effective teacher (evaluation tool score = 4.0 - 5.0) performs in a manner that consistently exhibits multiple strengths and that has a strong, positive impact on students and the school climate. The teacher serves as a model. Areas for professional growth are self-directed, based on refinement areas identified through the TAP rubric used by JHS. An Effective teacher

PERCEPTIONS OF TEACHER EVALUATION MODELS

(evaluation tool score = 2.26 – 3.99) has performance that typically exhibits multiple strengths and that favorably impacts students and the school climate. This teacher serves as a model in some areas. Areas for professional growth are self-directed, based on refinement areas identified through the TAP rubric. A teacher evaluated in the Improvement Necessary category (evaluation tool score = 1.6 – 2.25) has performance that typically exhibits few areas of strengths. Areas for professional growth will be identified, and a remediation plan will be developed. An Ineffective teacher (evaluation tool score = 1.0 – 1.59) performs in a manner that frequently exhibits weaknesses and that negatively impact students and the school climate. Areas for professional growth will be identified, and a remediation plan will be developed. Direct and immediate intervention is required by an Administrator.

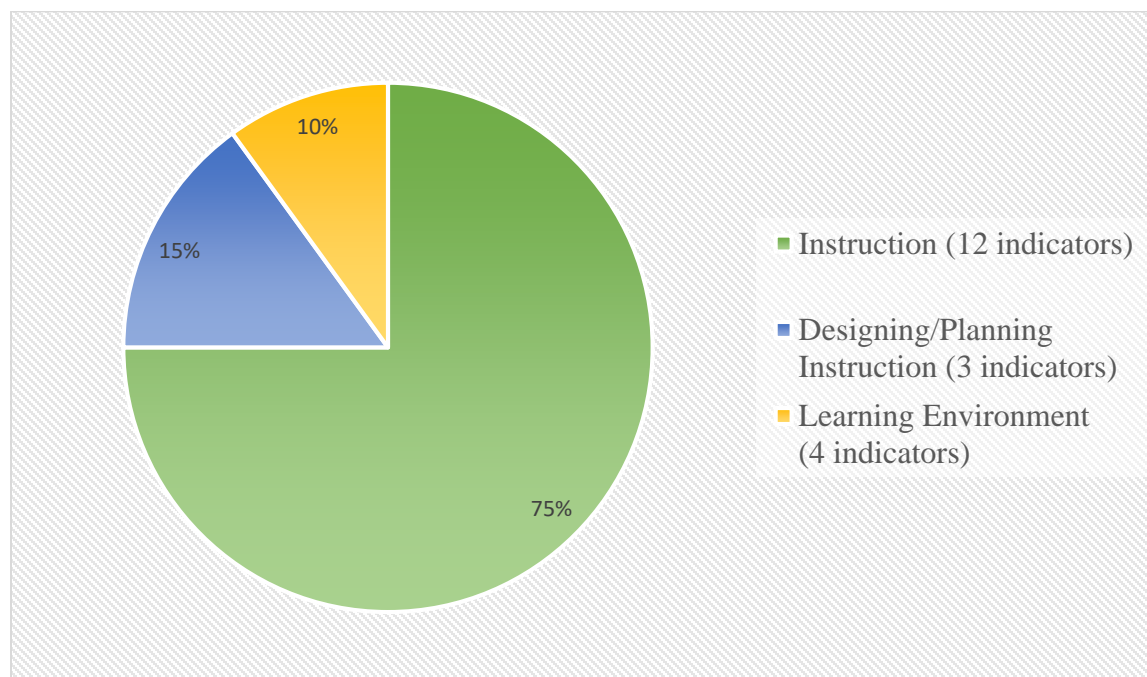
Procedurally, evaluators will evaluate (via observation criteria) non-administrative certificated personnel three different times during each school year, two announced and one unannounced evaluations. At JHS, the evaluation team includes Administrators and Master Teachers, and each staff member shall be evaluated by multiple evaluators, based on Freda's discretion.

The evaluator, following each formal announced or unannounced observation, will conduct a post-conference with the teacher who was evaluated. The evaluator will share with the teacher the evaluator's scores, and the teacher will share his or her self-rating scores during the post-conference; the evaluator's scores are the official scores. The evaluator will provide both one area of reinforcement and one area of refinement during the post conference.

Figure 1 illustrates how JHS weighs the three rubric domains.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Figure 1

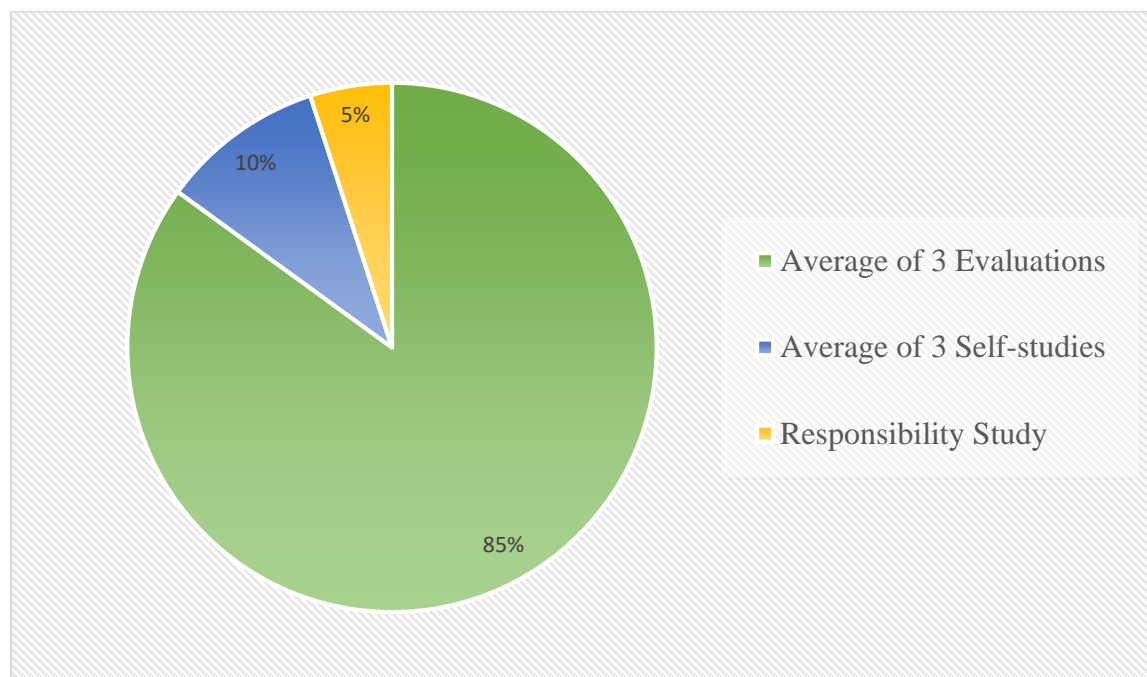
Weighting of TAP Rubric Domains at JHS

The CODE (Comprehensive Online Data Entry) system (software program) calculates final, integrated cumulative or summative performance scores as shown in Figure 2.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Figure 2

Percentages used to calculate cumulative or summative performance scores at JHS



Thus, as is evident, objective measures of student achievement and growth significantly informing the evaluation.

Jefferson High School: Participants

Freda worked at other schools prior to coming to JHS but previous to using the TAP model at JHS, she had evaluated teachers without using a specific model. Danielle's five years of teaching experience have all been at JHS. Danielle has been teaching various non-core classes in the same department her entire tenure. Evan's 17 years of teaching experience have also all been at JHS, and he has taught a broad range (introductory to AP and dual credit) of courses during his career. Evan has taught at the same grade level for the last few years.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Jefferson High School: Themes that Emerged from Research Question 1

The first set of themes emerged as a result of data collected from participants' responses to the first research question: What characteristics of the TAP teacher evaluation model do teachers and administrators perceive to be the most effective in promoting professional growth? Three themes emerged: training under TAP led to growth; TAP caused changes to instruction; and feedback from post-conferences and master and mentor teachers caused growth.

Theme 1: Training under TAP led to growth

First, a theme that training opportunities offered with JHS's TAP model led to professional growth emerged. First of all, all three participants described things before the TAP model was used. If you were a teacher before the TAP model was adopted, you met once with an administrator every other year or so and only to set an individual goal with few parameters. Evan said he had not had someone do a classroom evaluation for years prior to TAP being implemented. Danielle, who started the year TAP was adopted, knew that the new model was something very different than what had been done before. Yet, from the beginning, the evaluators (the five administrators and four master teachers) received training from NIET, and in turn all teachers received training. Freda made it clear that in-depth training occurred and that she believed it "served to give the feeling for those trained that [they] were all in this together."

JHS's TAP model includes a component of weekly cluster meetings to promote professional development. Freda described the groups as "15 to 20 teachers that would work together to gain and share insight and to answer questions." The cluster meetings are 50 minutes long, and the master teachers run the meetings. In the first year, the entire staff went through each of the 19 indicators one at a time in their clusters. Also, since then, all teachers review the

PERCEPTIONS OF TEACHER EVALUATION MODELS

TAP indicators at the start of each year (the first nine weeks), and teachers new to the district undergo additional, more intense training in after school sessions. Evan stated,

I believe the school did a good job of taking time to go over every indicator as a whole staff. This occurred in the first year and then for a cycle every week for the first nine weeks or so. Professional development then continues to be framed in the cluster weekly meetings. As the years have gone by, the teachers still get trained every year for a length of time with cluster meetings.

Danielle said of those cluster meetings:

We learned from each other what would good, best practices lessons look like, and broke that down in relationship to the indicators. We underlined, highlighted, and discussed the rubric and discussed activities that could be thrown in here and there. Both administrators and master teachers would do some walkthroughs to see it in practice and use the observations to help with the teaching in the cluster groups.

She concluded that she could see from those sessions “how the rubric could be used in my classroom to help make me a better teacher.” Further, in addition to the bare bones of the rubric, Danielle said that teachers are given a handbook that contains examples “and is a good resource if I feel the training is not enough” on a certain point.

After the refresher at the start of each year, the cluster meeting focus on all sorts of professional development topics. Freda said that “[p]rior to the TAP model being implemented, professional development was just something we did as a hit or miss from week to week.” But now, the cluster meetings are purposeful and planned. Even to the degree, as mentioned by Danielle that “sometimes the administration has given [the teachers] a survey to find out which indicators need a bit more training provided to [them] as a teaching staff.” Evan said that having

PERCEPTIONS OF TEACHER EVALUATION MODELS

the option in TAP to collaborate with fellow teachers is critical. He said that having conversations about teaching with colleagues causes professional development “because any time you have confident exchange and you can take away something that you can implement in your classroom growth is occurring.” Relatedly, Freda discussed the practice of doing walkthroughs and then having conversations with the master teachers so that they can provide instructional assistance in areas that are revealed as needing improvement. Both teachers mentioned the walkthroughs in positive terms.

Theme 2: TAP caused positive changes to instruction

The second theme that emerged was that the TAP model causes teachers and administrators to make positive changes to their instructional strategies. Both teachers refined their instruction, and Freda changed evaluation practices. Freda said she changed the way scripting happens with evaluations since using TAP and that she now “very much know[s] what to look for, which ultimately has helped [her] become a lot better at what I can say to teachers before and after the evaluation.” She added, “The TAP model made me much more intentional . . . with a clearer vision about classroom instruction; I’m less judgmental,” which has helped her in relating to teachers. She said that using the rubric “is a big advantage as now [she] can define what good instruction looks like.” Freda added that she “believe[s] the instructional indicators in the model do make sense and are clearly written to define focused instructional components” and that she “ha[s] gotten much smarter and better at the evaluation process as a whole.”

Danielle and Evan were very similar in their comments as related to their classrooms. Danielle stated, “I thought I was a good teacher to start, but the model made me more aware of things that I was not necessarily doing or could do better. The TAP model made me more purposeful in my actions and aided in my understanding of what I was doing in the classroom.”

PERCEPTIONS OF TEACHER EVALUATION MODELS

Likewise, Evan said, “I would say that I think I am a better teacher as a result of the TAP model.” Evan also mentioned being more reflective and understanding what he could do better, concluding that “[he]’d be lying if [he] said [he] did not get better as a teacher by utilizing the indicators on the rubric.”

Danielle became more purposeful in what occurred in the classroom to meet the instructional indicators for students. One particular example from Danielle was that she “changed the ways [she] purposefully plan[s her] lessons and [her] expectations for student learning to be evident and shared.” Once again, Evan agreed. He said that “one of the big things TAP does is it makes the teacher more reflective if they are really taking what they are going seriously.” Evan uses the rubric to validate what he is doing well. He said, “It validated the things I was doing right or well, particularly in my classroom management, knowledge of students, and content structure.” In fact, Evan believes that TAP helped him “understand the importance of knowing [his] students” since, in his opinion, “having those relationships and knowing that is the bedrock of any successful classroom experience.” The TAP model has made Evan more reflective and purposeful. Evan said something that arguably best sums up this first theme: “I think people even who started begrudgingly with the TAP evaluation model would say they are a better teacher because of this model . . . because teachers are now doing more thinking about their teaching as a result of the rubric.”

Theme 3: Feedback from post-conferences and master and mentor teachers caused growth

The third theme was that the TAP model’s use of feedback is considered by all participants to be a very real positive that facilitated growth. The first aspect of feedback discussed by all three participants had to do with beliefs that growth came about from post-conference dialogue and feedback. All three participants agreed about the importance of

PERCEPTIONS OF TEACHER EVALUATION MODELS

feedback and dialogue in pre-and post-conferences. In these sessions, teachers get to do most of the talking and administrators ask needed directional questions in order to see areas for refinement and reinforcement. The conferences allow for discussion centered on indicators listed in the rubric. Danielle talked about an example where she changed her teaching in the classroom based on feedback provided by the evaluator early in the TAP model implementation cycle:

One example of growth was my evaluator told me in my post conference that they wanted me to refine my academic feedback [a TAP model indicator]. Not necessarily how I was giving the feedback, but how the students could give academic feedback as well. I was not happy with the feedback at first, but after I tried it, the students met the feedback expectations better than I would have imagined. This gave me the opportunity to focus on the students who needed the greatest amount of help and assistance.

Danielle liked “getting little suggestions or nuggets of knowledge from the evaluators who have more experience and knowledge at this point in their career than [she does].” Freda stated, “The comfort level of working through conferences together to discuss needs and components of the evaluations has brought alignment with teachers and evaluators.” Evan agreed with this theme, adding that collaboration after the conferences with fellow teachers is also a good element for growth. This allows further instructional discussions to occur and work to happen on curriculum changes, which is very beneficial when looking at instruction. Evan described being given a score in the low- to middle-3s the first evaluations, but raising that to the highly effective range (4.0-5.0) as a result of the feedback he was given.

A recent change to JHS’s TAP model has made pre-conferences optional if the teacher chooses (and unannounced observations necessarily cannot include a pre-conference), yet Freda

PERCEPTIONS OF TEACHER EVALUATION MODELS

believes that taking pressure off those teachers who choose not to have a pre-conference meeting has rewarded teachers a bit and has not taken away from the positive growth that feedback causes. Interestingly, Danielle said, “Most of us teachers still go ahead and do the pre-conferences because we think it helps your evaluation and your evaluator.”

Secondly, all three participants agreed that master and mentor teachers were used effectively under the TAP model to give feedback. Master teachers function in a unique manner relative to the traditional teacher in a TAP school. They do not teach classes during the school day, but instead design and implement teacher professional development in the cluster groups and assist/guide colleagues in accessing or selecting appropriate research based TAP model strategies. Master teachers serve to build instructional capacity in all teachers through evaluations and coaching. According to Freda, in addition to master teachers, JHS introduced mentor teachers to the evaluation model last year. This involved picking a teacher from every department to be a mentor teacher. The mentor teachers meet with the master teachers once a month and have a common planning time to meet and discuss staff needs. The mentor teachers serve as the go-between for evaluators and evaluatees, and are often used as examples or as short presenters for the whole faculty at cluster meetings. Freda expressly said the master teachers’ “job is to serve the teachers” and that she believes the master and mentor teachers “help with staff buy-in” and, from a “visual sense, provide the opportunity to see and talk about classroom strategy and improvement.”

Danielle recounted being willing to talk with a master teacher about frustrations that she was unwilling to talk about in her full cluster group. She said she “built trust and confidence in [her]self through the developing relationship with the master teachers” causing “her receptiveness to feedback to gr[o]w and bec[o]me much more usable.” Danielle also said that

PERCEPTIONS OF TEACHER EVALUATION MODELS

she liked being evaluated by these different teachers along with the administrators “because you get a different lens each time”; although she did acknowledge a potential for “rater-reliability issues to occur.” Evan said that the idea of mentor teachers “has worked for [him] because [he] has been lucky enough to have good ones that [he] respected and trusted early in [his] career and that is the role that current master and mentor teachers are supposed to serve.” Evan described a process where, for teachers who need it, “the master and mentor come in to provide resources, talk the teacher through some best practice strategies, and share lesson plans to analyze where improvement could take place.”

Jefferson High School: Themes that Emerged from Research Question 2

The second set of themes emerged as a result of data collected from participants’ responses to the second research question: What characteristics of the TAP teacher evaluation model do teachers and administrators perceive to be the most effective in promoting student achievement? Two themes emerged: use of student data is lacking under TAP; and TAP created a more student-centered classroom.

Theme 1: Lack of student data

Evaluation of Research Question 2 was influenced by the consistent theme that the TAP model as used at JHS does not require teachers to track, use, or report data on student achievement. All three participants agreed about this, and the information provided from all three sounded nearly the same. While data is apparently looked at occasionally in a few circumstances (e.g., department meetings), teachers were not responsible to provide and administrators did not ask for student data. Danielle said:

I have not been asked to provide student data at all. For the most part, it has been more anecdotal and individualistic in nature and for my purposes only. Nothing has to be

PERCEPTIONS OF TEACHER EVALUATION MODELS

turned in as a teacher to be part of a final percentage of your yearly review. So, overall the TAP model score for the year has no component of student data included in the final tabulation; even through the post conference there really is no discussion at all about student data.

Evan said nearly the same thing, although he did describe some cluster meetings where “Teacher 101 type of data stuff” (“what are things we did well and what are things we have to touch up on” as a teacher) was looked at. He also acknowledged a program used in math and English courses that generated a monthly score on assessment reports, but described a program that does not work consistently and does not have good support from the tech department, which he said resulted in teachers being precluded from using it effectively. Freda said that “teachers do not want to get bogged down in cluster meetings with data.” She concluded with “teachers do not have to do any data tracking at all for the year under the TAP model.”

Theme 2: TAP created a more student-centered classroom

Despite the lack of data discussed above, a second theme emerged that TAP created a more student-centered classroom. In fact, Freda saw this as the biggest change brought about by TAP. She said students used to be sitting quietly at their desks doing work individually or while teachers were lecturing. Now, she routinely sees “student-to-student interaction, . . . students asking questions, [and] students learning from each other. . . . The students are exploring; they are curious.” She attributes this to the passive vs. active learning indicator on the TAP rubric “pushing this to happen.” Freda explained further that if teachers want to raise their score from a three to a five, they “must have a deeper and richer classroom experience. In this instance, are students asking higher order questions? And, for that to happen, [the teacher] has to let them talk and have some free rein, and I have witnessed that happening.”

PERCEPTIONS OF TEACHER EVALUATION MODELS

According to Danielle and Evan, it is being seen because it is, in fact, happening in their classrooms. Danielle said:

I do less “sit and get” and more active participation now than I did at the beginning of the TAP model. Allowing time for students to give each other feedback was a newer strategy that has become a key piece to the culture of my classroom. I have always been a hands-on teacher first and foremost, but the model helped me incorporate different strategies allowing me to hone my skills of being an active communicator with my students.

She talked further about having students who now are challenged to reach deeper and engage in higher order thinking. Evan said virtually the same thing:

My predominant instructional strategy is trying to get the kids doing more and me as the teacher talking less. I know that most students probably absorb more when they are doing something, whether it is a kinesthetic activity or even just something working in a group where I am assigning roles and saying get this done and then share back the results. That is my biggest strategy using grouping and getting them active and working.

Grouping of students has been an area of refinement written in the past TAP evaluations.

Evan also talked about how he feels that his focus on the “last bullet point in the rubric descriptors” involving students leading and taking ownership for their learning has garnered results. As he explained:

My students generate questions, students given specific feedback to one another; they show they have mastered the objective. When I have felt I meet those expectations with my lesson or unit, then student achievement had a chance to improve. . . . Reaching that characteristic of the rubric and model has been my goal and what I have seen make the biggest difference in student achievement.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Jefferson High School: Themes that Emerged from Research Question 3

The third set of themes emerged as a result of data collected from participants' responses to the third research question: In what ways do teacher and administrator perceptions of the TAP evaluation model influence the evaluation process and educational outcomes? Two themes emerged: teachers' fear of the model was reduced through use, which positively affected the use of TAP; and there is a certain degree of buy-in to the TAP model that influences teachers' desires to use it for improvement.

Theme 1: Teachers have gotten more comfortable with TAP

The first theme that emerged was that teachers' initial fear of the TAP evaluation model was reduced through use, and this positively affected the use of TAP. Basically, as teachers got more comfortable with TAP, they perceived it more positively and were more willing to engage with the model. Freda spoke to this directly, and while the teachers were not so direct there were enough comments throughout their answers for it to emerge as a theme. Freda said that the staff was "a lot less threatened by this model than they were initially." She talked about a veteran teacher who was new to JHS who was quite nervous about being evaluated, and then she said "I would be too even though I have been in education for 34 years." Freda said she believed that the teachers do not hate the evaluations like she thought they might when TAP was first implemented and that TAP "is not looked upon as a 'gotcha model' by the staff." She saw their increased comfort level as a contributing factor in the overall success of the model. Freda concluded that "[b]ecause the staff feels the collaborative efforts, because it has not affected their salary, and because even the hard core teachers say they are a better teacher than five years ago, the TAP model works for us." Relatedly, Danielle spoke about teachers starting out fearful when the model was first implemented but how she then saw that change, and Evan talked about a

PERCEPTIONS OF TEACHER EVALUATION MODELS

level of angst that was present when the early cluster meetings focused so heavily on the TAP rubric, which resulted in “professional development being stifled” but then made it clear that this perception has changed for the better.

Theme 2: Buy-in needed to use TAP to improve

The second theme related to research question 3 is the perception that there is a certain degree of buy-in to the TAP model that influences teachers’ desires to use it for improvement. The three participants were largely favorable about the TAP model, feeling that it had contributed to both professional growth and increased student achievement, and this perception meant that they were using it. Danielle indicated about the model:

I think the model completely aligns with good lessons and a well-run classroom. Every indicator in the model is lined up with best teaching practices and what is best for students to get them to grow. The descriptors have helped because they give specific detailed items to incorporate into the lesson. For example, the model helps me think on my planning am I making powerful connections by teaching my content in different ways to different kids and their learning styles.

Evan was perhaps not as enthusiastic, but he still admitted:

Ideally, the rubric domains and indicators are aligned with one another, meaning that if a teacher excels in one part of the rubric, they will most likely excel in other parts. For instance, a teacher’s knowledge of students could impact lesson structure and pacing, which then impacts measuring student mastery. All three of these indicators seem to address both academic and classroom management/structure in a way that illustrated the symbiosis among the indicators.

PERCEPTIONS OF TEACHER EVALUATION MODELS

He said, “The model helps me to be more reflective in my craft, making me assess my decisions and actions in the classroom more deliberately.” Freda similarly stated that she believed that “the rubric places emphasis through the indicators on the right ways as to lay out expectations for the staff so all of us know collectively what is expected of each other in all classrooms across the school.”

Also, while all three participants referenced other staff members who were not so favorable about TAP, they all seemed to generally feel that most teachers had at least “accepted it” and were using it with favorable results. Freda said that all teachers at least see it as “a necessary evil,” and Danielle said that there were teachers who saw the model as a checklist of items to complete, “[b]ut overall, I think the teachers realize that the TAP evaluation model is here to stay and have accepted it and gotten more comfortable with the process.”

Similarly, Evan said:

Most teachers have accepted the model after several years of implementation. They understand that while this is a comprehensive and expensive model, there are others that may not have the quality of TAP. Teachers are not enthusiastic about TAP, but they accept it as a whole and that it has allowed us to satisfy the law’s and DOE’s expectations.

Thus, it seemed clear that the success of the models was influenced by the generally positive perceptions held by the JHS faculty.

Jefferson High School: Summary

In this chapter, the findings from Jefferson High School (Case Study #2) were reported in order to answer the specific aspect of the overarching research question relevant to JHS: from the viewpoints of both teachers and administrators, does the TAP teacher evaluation model actually,

PERCEPTIONS OF TEACHER EVALUATION MODELS

in practice, cause teacher growth and improve classroom instruction? The findings of the three interviews were coded and themes were determined as they applied to the study's three research questions. With regard to research question #1 pertaining to characteristics of the TAP model that the participants identified as most effective in promoting professional growth, three themes emerged. First, the participants agreed that the training that was provided under TAP was an important aspect leading to growth because it ensured the whole staff working together to understand how to implement best practices and fostered a spirit of collaboration. Second, the TAP model inspired the participants to be much more intentional with how they approached the rubric items. Teachers and evaluators knew exactly what evaluations would involve, and thus what happened in the classroom (both teaching and evaluating) was more planned and developed with best practices as the guiding force. Lastly, it was clear that feedback under TAP was a valuable aspect of the model. Most teachers were taking advantage of the pre-conferences and the mandatory post-conferences were seen as very important learning mechanisms. Further, the model's master and mentor teacher approach was seen as directly contributing to growth.

With regard to research question #2 in which the participants were asked to identify the characteristics of the TAP model that were the most effective in promoting student achievement, two themes emerged. First, it was clear that the lack of required use of student data affected the ability to accurately measure specific student achievement gains. While growth overall was evident, the participants indicated that use of data would make the nuances available and this would be helpful. Second, TAP clearly created a more student-centered classroom. The participants could easily think of ways that the classroom component had very intentionally changed and improved to benefit student achievement under TAP.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Last, with regard to research question #3, which explored the ways that participant perceptions of the TAP model influenced the evaluation process and educational outcomes, two themes emerged. First, it was apparent that teachers' fear of the evaluation model was reduced through use. TAP was not perceived as a "gotcha system," and this comfort level positively affected the use of TAP by all involved. Finally, the last theme was that there is a certain degree of buy-in to the TAP model that influenced most teachers' desires to use it for improvement. While there were certainly teachers (usually those with the most experience) who may have preferred the pre-TAP days, even those teachers had accepted that the evaluation system was there to stay and had gotten on board with using its positive features. Thus, in all, the positive perceptions led to use which allowed the system to work.

PERCEPTIONS OF TEACHER EVALUATION MODELS

CHAPTER SIX: FINDINGS – CASE STUDY #3

The purpose of this qualitative case study is to analyze three High-Stakes Accountability Models—RISE Evaluation and Development Systems (RISE), Teacher Advancement Program (TAP), and Peer Assistance and Review (PAR)—and this chapter reports on Lincoln High School, which is part of a school district using the PAR evaluation system. As previously explained, data was collected on teacher and administrator perceptions concerning: (a) the general methodology or structure of the three evaluation models; (b) the specific practices for gathering and compiling the raw data in each of the three models (i.e., the forms used); and (c) the potential impact of each model on teacher effectiveness, student achievement, and professional growth.

At Lincoln High School, I interviewed Ian, an administrator with the most knowledge about and experience with the evaluation model used by the school; Gary, a teacher with five or fewer years of teaching experience; and Hillary, a teacher with six to twenty years of teaching experience. This chapter will describe both the school profile and the evaluation model used at the school, background on the participants, and the themes that emerged from the interviews.

Lincoln High School: Profile

Lincoln High School serves 200 students in grades 9 through 12. It is located near a small, rural community. The junior high and senior high share a building. LHS draws students from several small surrounding towns as well as some tuition students from a larger neighboring school district. All grades are on the same bell schedule and share several classrooms and the cafeteria. Generally, students are in classes with their same aged peers. The demographic breakdown of LHS is shown in the table below.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Table 6

Demographic Breakdown of LHS

LHS student category	Percentage
Special Education	10
Free and reduced lunch	38
White	96.2
Hispanic	2.3
Multi-racial	.8
American Indiana	.4
Asian	.4

During my interview with the LHS administrator, Ian, he described the following about LHS: a 1:1 school, Lincoln students each have a computer that they use both at school and at home. Though each student has a device, two labs are available as well for classroom use. The school employs two technology staff members. Collaborative integration of technology is a curriculum focal point. Twenty-seven full time teachers are employed by the LHS. In addition, the special education teachers work with four paraprofessionals. The administrative staff consists of a principal, an assistant principal/athletic director, a guidance counselor, a media specialist, two administrative assistants, a maintenance director, four cafeteria workers and three custodians.

Lincoln High School: PAR Evaluation Model

As long as the three participants utilized in this study can remember, there has always been an evaluation system in place, even prior to P.L. 90, which required evaluations to take place in all school corporations. The evaluation system prior to the law changing was very basic, and very little information was shared in the final evaluation. According to one of the teachers interviewed, “I felt the evaluations were done with the overall interpretation left up to the administrator. It was very broad and open ended. Compared to now, where it is very spelled out

PERCEPTIONS OF TEACHER EVALUATION MODELS

and prescriptive in nature.” When the law came into effect, LHS was picked by IDOE as a pilot school for the RISE model. According to the same teacher, “It was awful. We went from an evaluation document that had very little information given to the staff in terms of expectations to an entire RISE model binder that was just pages and pages. It was overwhelming.” When the school district had a chance to make a different decision after that pilot year, the school corporation went quickly away from the RISE model to the PAR model and has remained there ever since.

At LHS, the PAR evaluation has three components. The first is Classroom Observations, Part A; the second is Student Growth and Achievement Data, Part B; and the third is Technology Instructional Application, Part C. Parts A and B are combined for the teacher evaluation score. Points earned in Part C are added to the evaluation score. The Part A rubric for the PAR evaluation consists of two domains. Domain 1 is teaching procedures and Domain 2 is classroom management. Under Domain 1 there is a planning component, a delivering instruction component, and a motivation component. Domain 2 looks at best practices to ensure that the classroom is both orderly and an engaging learning environment. In addition, evaluators look for behavior and conduct being clear and consistently modeled and enforced.

Part B is where teachers will measure growth or achievement for all students who are assigned to them for a major part of an instructional period. Teachers in areas or grade levels where state-wide assessments are administered must use data from these tests. For the other categories of teachers, the growth measurement will be a pre-test/post-test assessment. Pre-tests will be given within the first six weeks preferably. Post tests will be given at the end of the class.

The PAR evaluation model for LHS is in handbook form and put on-line for all teachers to have access to anytime. Part C, or the bonus part, does not count against teachers, but can

PERCEPTIONS OF TEACHER EVALUATION MODELS

only help their overall grade. For example, according to Ian “the high school has invested a lot of money in technology, and we want to encourage staff members to use it in the classroom. So teachers can earn bonus points that are added into the total.” LHS believes that effective use of technology in instruction is a priority of the corporation and an important component of preparing students for success in the twenty-first century. The model is set up with the following percentages assigned for your summative total rating for the year. Sixty percent is based on the teacher’s final of the four observations, and 40% is on the data portion designed for growth. The bonus area can help improve a teacher’s total by 5% if a teacher chooses to try and obtain bonus points.

The PAR evaluation model has four types of classroom observations. First, there are two planned observations, lasting the whole class period. Second, there are two unplanned observations. Third, there is also an opportunity for a peer observation. Peer observations are intended to be professional development opportunities and cannot be factored into the final evaluation score. The peer teachers may discuss the planning and preparation for instruction in advance with the teacher so that the teacher knows the date and time of the observation. After each observation there is a post conference. The post observation conference must take place within 5 days of the classroom observation. The fourth type of observations are walk-throughs. Walk-throughs are not scheduled in advance; they are done as opportunity allows. In addition, walk-throughs were not intended to be evaluative. In fact, data collected from walk-throughs may not be considered evaluative without conferencing with that staff member.

Also, within the PAR model there is a governing board of review for any teacher who falls in the category of Needs Improvement or Ineffective. The board reviews the evaluation process; ensures that the assistance was completed in a satisfactory manner; receives reports

PERCEPTIONS OF TEACHER EVALUATION MODELS

from the building principal; and decided to retain the teacher in assistance, release the teacher from assistance, or recommend for cancellation.

Lincoln High School: Participants

Ian was a teacher at two different levels for eight years before becoming an administrator, and LHS is his second administrative job. Ian has only evaluated using PAR at LHS. Gary taught at two other school corporations prior to his four years at LHS. He is also a member of the coaching staff. Gary teaches freshman and sophomores in mostly middle-to low-level classes. Hillary's 16 years of experience have all been at LHS, where she has taught in two different departments and is involved in many of the school activities. Hillary has taught a range of classes from introductory to Advanced Placement classes.

Lincoln High School: Themes that Emerged from Research Question 1

The first set of themes emerged as a result of data collected from participants' responses to the first research question: What characteristics of the PAR teacher evaluation model do teachers and administrators perceive to be the most effective in promoting professional growth? Four themes emerged: positive changes to instruction occurred under PAR; feedback under PAR promoted growth; training facilitated growth; and PAR provides common and clear language for shared expectations.

Because LHS was different from WHS and JHS in that it has only been using the PAR model for the last four years and the participants had used another model (the RISE model) prior to that, some responses throughout the coding and, thus, the themes refer to both models. I took care to make this distinction clear.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Theme 1: Positive changes to instruction

The first theme was that the PAR model caused the participants to improve their methods of instruction. Hillary and Ian were very clear about this. For example, Hillary said, “Under the PAR model, I have grown as a teacher in ways that I never thought I would, simply because the rubric categories make sense and I understand how to accomplish them and how to get better as well.” Gary agreed generally, saying things like he grew as a teacher because “[t]he PAR observation rubric categories are all things that make sense to use as a collaborative staff evaluation tool.” Ian saw the PAR model as involving himself in his teachers’ growth. He said, “With the PAR model, I am building individual teachers up throughout the year. I feel that it is just as much on me as it is on the teachers who by the end of the year have grown and improved.”

When asked about specifics, Gary indicated, “The model has made me think about items when I am planning my lessons and how that ties to a good lesson. I now plan as if the administrators were going to walk into my class every day.” Hillary agreed stating, “The planning part of the rubric has helped staff with their lesson plans, which in turn has helped make a better well-run classroom.” Ian agreed in his thought that the PAR model was improving instruction because it caused teachers to create plans and ideas for a more successful classroom. Relatedly, Ian said, “It has improved my ability to look for teachers giving the students the opportunity to apply what they are learning.” One example, driven from the model has been the incorporation of bell ringer and exit ticket strategies for checks for understanding in the classrooms. This is a specific category on the rubric and under PAR teachers were changing what they were doing and using these strategies.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Theme 2: Feedback under PAR promoted growth

The second theme centered on the concept of feedback. Using the PAR model requires post-conferences to take place after each evaluation. All three participants indicated that as they have continued to get comfortable with PAR, having the requirement to sit down face-to-face to have professional conversations with every observation has provided growth opportunities for the teacher and administrative staff. Hillary said it this way:

Our administrators now give excellent feedback and push even me who has taught for 16 years to get better in my classroom instruction. I want that challenge I want that feedback from my evaluator. If you had administrators that always said everything looks good or great work without pinpointing areas, I think the model would not produce better teachers as easily.

Because of the required post conference conversations, the evaluation feedback is not finalized until after the post conference. Hillary indicated, “The post conference is a lot more effective because one, we never used to have anything after an evaluation and because, two, the teachers and administrators understand the specifics of the model more clearly when they go over it together.” Gary said:

The post conferences allow my administrator and I to sit down together to go over the evaluation. The administrator can give tangible explanations for how I was rated, offer suggestions for growth, and provide feedback and direction on all areas. This time also allows me to ask questions, seek clarity and wisdom, and to explain myself if necessary.

Ian said, “That interaction time one-on-one has helped me establish trust, build cooperative relationships, and also provides me an opportunity to learn from each other through the opportunity to solicit feedback for reflection.” He also said:

PERCEPTIONS OF TEACHER EVALUATION MODELS

Feelings are now that teachers welcome feedback because they were included in the implementation process. I know this to be true because staff was much more hesitant for any feedback or growth requests using the prior RISE model where fear and angst were the words of the day.

Gary affirmed this lack of fear: “No teachers are scared to ask their administrator to come into the classroom when they are trying a new lesson and simply want feedback.” Gary also mentioned the immediacy of the feedback as a positive: “[The feedback] is nice, . . . [and] that it’s fairly instant feedback, not sitting there wondering if they thought that went good, because I did.” And, Hillary added to this, saying:

I feel the administrators are very honest with us, and if there seems to be some indication of problems with instruction or data, administration will be working with those teachers well in advance. You get to that second evaluation; you can begin to get peer assistance right away to seek our help on needed growth areas. The hope is by the time the teacher is at the fourth evaluation, they’re much improved in the growth areas instead of looking over your shoulder wondering if you’re going to lose your job because nobody helped you between evaluations one and four.

Ian said this about feedback:

it provid[es] assistance to help improve others. . . . I love being able to say to a teacher that if we do this, or use some of the best practices from the mentor teacher, we are going to get better at teaching together because I feel I can provide resources. Because I am sharing where I would like to see the teacher grow, I feel the evaluation model has stretched me to be more directive in my expectations of my staff. I have been stretched to be a better communicator as well.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Related to the conference feedback, is the peer assistance aspect of the PAR model. The PAR model is set up with peer assistance as a key component of its structure and for LHS, that peer assistance comes in the form of mentor teachers. Gary said:

Having mentor teachers is a huge selling point with this model. It is sometimes much easier to hear the tough conversation stuff from the person you have a side by side working relationship with. Mentor feedback is often deeper, richer, and for me early in my career more meaningful.

Gary gave a specific example of a mentor teacher suggesting in post-evaluation feedback that he give seating charts a try even though he “hated seating charts.” Gary believed his justification for not using seating charts, but they talked about it, and Gary respected the mentor’s experience, so he gave it a try and has continued to use the modified seating chart that was suggested with good results.

Hillary also spoke highly of the mentor teachers and their ability to assist in growth of the teachers. She said that under PAR, having the additional sources of feedback takes away the excuse for teachers that they did not know there was an issue. Through the mentor teachers, the administrator can work with the teachers well in advance of the final evaluation to try and correct problems or issues. Gary affirmed this notion, saying, “Mentor teachers have the experience, so I tend to listen very carefully. Because of their involvement, the PAR model has given me some good ideas to think about.” Gary went on to state:

having the mentor teachers absolutely helps the most. Having someone who has taught anything for several years, obviously they are pretty good at what they do at that point. It is good to have an experienced teacher down the hall to walk down to and ask questions, seek guidance, and bounce ideas off of them.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Ian said the mentor teachers were given a checklist of what they needed to help new teachers with, and they were expected to go through and make sure that everything got checked off. The mentor teachers are not seen as being helpful to only the teachers, however. Ian said:

This model helped me see that evaluations are not just me by myself. When I get into a classroom and I see we need to work on classroom management for example, I know in the PAR model I can pull in resources, i.e. two expert teachers in that area, to help work together with the teacher who needs assistance. This model has promoted assistance from mentor teachers and not just me, allowing me to not have to feel like I must be the expert in every area.

In addition to mentor teachers affirmatively reaching out, any teacher can request to see other teachers in action. “This is just an expectation of the PAR model that can and is being taken advantage of by the staff,” indicated Ian.

The peer review component also includes the peer review panel. Both Hillary and Ian talked about their role on the peer review panel as a characteristic of growth for both of them as well as those teachers served. The peer review panel is put into action when there is a teacher in need and all other support has been utilized. Hillary said:

Being on the panel has increased my professional relationships with my fellow co-workers. When they see you involved in an additional layer of efforts to support one of our own, the phrase I often hear is this is a school that does everything they can to support teachers.

From an administrator viewpoint:

It is advantageous as the leader of the building to see situations work them self out both positively and negatively through the peer review panel. If it works out positive, the

PERCEPTIONS OF TEACHER EVALUATION MODELS

teacher has improved their craft, and we continue our working relationship. If it turns out negative, then the staff understands I tried every possible means to provide assistance, and the teacher simply did not get the job done. It also serves to show that I will take staff to a peer review panel if they are not getting the job done and that we mean business in the education of students. In both cases, it makes me feel good that the process works under the PAR model (Ian).

Theme 3: Training facilitated growth

The third theme that emerged had to do with the training that happened with the PAR model. As mentioned, JHS's PAR model was developed through the work of the PAR committee. That committee also developed the initial trainings related to each aspect of the model's rubric. With that in mind, Hillary (who was on the committee from its inception) said:

Basically the PAR committee continues to lead the training and to lead any potential changes that need to be made in ongoing implementation. The PAR committee does training and rubric tweaks over the summer for new teachers and administrators.

According to Gary, "[a]t the beginning of the school year at the opening staff meeting and through mentor teachers, the PAR model retraining is done, and time is spent going over questions or concerns." Both teachers also indicated that as the PAR model continues to be in use, there needs to be the same emphasis on training and retraining of the model as the staff gets further away from it being first implemented or the training they once had would cease to carry over.

Hillary said that at LHS, PAR training has really got them trying to focus on positive change. She said:

PERCEPTIONS OF TEACHER EVALUATION MODELS

This model is definitely about growth, and a teacher does not have to be perfect in every category right from the start. The RISE model felt like those were standards you would never reach and not focused on the growth component for teachers. Teachers have bought into this PAR model because they think it is attainable.

Hillary felt that training showed them that the PAR evaluation model really modeled what should be happening in the classrooms.

Ian found that mentor teachers help train or re-train newer teachers on PAR; also, “the mentor teacher “helps provide assistance with grade protocol, technology, Schoology, and PIVOT training.” Further, for Ian, the mentor teachers have stretched his thinking about consistent training. “The PAR model has stretched me in my growth to become more directive in my communication of my expectations for the staff. I never want to put the mentor teachers in the spot of not knowing what exactly to say or do.”

Theme 4: PAR provides common and clear language for shared expectations

The fourth theme was that the PAR model provides common and clear language for shared expectations. The common language used by the three participants made it evident that they had the same expectations that the evaluation rubric is cumulative in nature, measuring growth throughout the school year, and that final classroom evaluation ratings are given on the last planned observation of the school year. Hillary specified that knowing exactly how the system works “helps build the capacities and strength of a teacher throughout the school year.” She added that under PAR, she has been able to focus on “good teaching every period” and not the “nit-picky items” on the RISE model, like being worried about getting “yelled at for not having our standards on the board.” She felt that the PAR standards “were more focused on what and how the material is being presented in the classroom.” Ian said that he “is much more

PERCEPTIONS OF TEACHER EVALUATION MODELS

clear on what [he] is looking for when going into the classroom to observe” and that the “clarity is provided by the rubric set up under PAR.” The rubric, he feels, helps him to help the staff focus on “bell-to-bell instruction” and helps him know when teachers are doing something all the time versus just for the observation.”

The number of evaluations and the length of time spent in the classroom promotes a common understanding between evaluator and evaluate. Ian said, “This provides accountability for my teachers. Having a constant scope and sequence for every class and every teacher keeps an understood level playing field. This allowed students to have a better teaching experience from which to learn.” Further, the model gives teachers “a set group of categories that they need to try to hit during each lesson,” and Ian feels this clarifies the difference between “accountability and micromanaging.”

Additionally, all three appreciated having a consistent location for the evaluation database and felt the staff was comfortable with the PIVOT layout and the ability to give or get instant feedback. Gary stressed this point, “Evaluations are done on-line through PIVOT. We can see our past evaluations, we can see our now evaluations with fairly instant feedback, and what is still left to come. It is really nice because I can see what was written and come up with my questions even before we start our post conference meeting.” Ian further reiterated, “The common and clear language has allowed me to be clearer and focused as a classroom observer. It has helped me pinpoint good teaching techniques based on the rubric.” Similarly, if what was observed in an early evaluation was not good, then this model has checks and balances to make sure it will not ruin the teachers’ grade for the year.” From this idea, the three pointed out that they have all grown because of the common understanding of expectations.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Lincoln High School: Themes that Emerged from Research Question 2

The second set of themes emerged as a result of data collected from participants' responses to the second research question: What characteristics of the PAR teacher evaluation model do teachers and administrators perceive to be the most effective in promoting student achievement? Two themes emerged: PAR caused improvement in teachers' instructional strategies; and PAR required the use of student achievement data and this helped with student achievement.

Theme 1: Improved instructional strategies

Regarding the first theme concerning instructional strategies, all three of the participants reported that "yes," strategies have changed since the implementation of the model to help students achieve. However, because Gary and Ian both said PAR had resulted in a change yet predominantly discussed their original teaching strategy, I had to closely review what they told me and rely on my field notes pertaining to my perceptions of their overall outlooks. Both Gary and Ian talked about their predominant instructional strategy as "I Do, We Do, and You Do." Both indicated that it is a primary teaching strategy for developing guided instruction and that it has been since they started teaching or before thinking about the PAR model. Yet, it was very clear that Gary and Ian, as well as Hillary, had in fact changed their instructional techniques under PAR. Gary stated:

An example of this [change] would be student engagement in the classroom and a check for understanding. . . . For me personally I know the PAR rubric has helped me try to increase my engagement. I am much more conscious about having all students engaged, and I have a check sheet that I keep on my clipboard to make sure I am not leaving anyone out. That change has come directly from the PAR model component. Also, I

PERCEPTIONS OF TEACHER EVALUATION MODELS

have changed from a lecturing style to more of a discussion. This model has also helped me develop skills to be even more hands on with the activities with the students.

Hillary similarly gave an example of a specific change that directly impacted student engagement and learning. She indicated that prior to the model, her main instructional strategy was to lecture the material to the students and then give them time to show knowledge of the material through homework, quizzes, and tests. Now, however, Hillary is very purposeful about increased student engagement. She said:

The model led us to become more concerned with what our students are doing in the classroom. Student engagement and checks for understanding are prevalent in my classroom now because of the categories on the evaluation rubric. I have changed because of the model. I am much more student centered and have changed from a lecturing style to more of a discussion style. Quite frankly, becoming much more hands on in my instructional approach has also taken place because of PAR.

Similar to Gary, Hillary now has a specific method for bringing about student involvement and engagement. She uses a program on her smart board that randomly generates which students will be called upon. This ensures she is picking all students to engage with class discussions. Hillary also, “changed the way [she] ha[s] students write out their answers and show work. This has helped change the way [she] checks for understanding throughout the class period.” She said, “I probably would not have changed my practices without seeing the category on the PAR model.” Hillary believes that her students are growing and PAR lets her “be professional in [her] determination of where and how kids can grow based on their knowledge and level of ability.”

Ian confirms that he has seen the move from the “whole group lecture style of teaching” and has seen the results of focusing on “guided instruction and best practices for student

PERCEPTIONS OF TEACHER EVALUATION MODELS

growth.” He is also seeing better planned lessons because teachers are using an on-line planning book that see what works and what does not and more comfortably plan different activities for the classroom. Thus, despite the growth not being as fast as he’d like, he was clear that he believed it had occurred.

Relatedly, the teachers felt that PAR emphasized instruction and instructional strategies rather than micromanaging classroom management. Gary talked generally about how an evaluator without the rubric just looked at the outside, at how things seemed to be going, but under PAR they can look for specific things and use the data to see that growth is actually happening. Hillary also said that “with PAR we are getting back to what and how material is being presented and taught in the classroom,” adding that “[t]he PAR model has led us to be more concerned with what our students are doing in the classroom, and the RISE model was much more teacher centered.” Ian talked about meeting the needs of the individual teacher in the classroom rather than looking at the evaluation as involving whether a checklist was mastered. He said that he has heard and seen from teachers that RISE was “more trivial work on their part and the focus was not on the important areas. So everything we develop in this model ties to exactly what they are doing in the classroom. So [what they are doing] is working hand in hand with the PAR model. . . . The PAR model has helped them because they have more interest in getting better under this model.”

Theme 2: PAR required use of student achievement data increased student achievement

The positive use of data was the second theme related to Research Question 2. As Ian said, “A characteristic about the PAR model that is good is the data component. The pre-test, post-test option lets teachers see the growth that has occurred throughout the course of the year.” He went on to say that:

PERCEPTIONS OF TEACHER EVALUATION MODELS

[t]he model has required data must be used to drive instruction and be shown in the evaluation process. Teachers are expected to keep track of the data and report it at the appropriate time. Data growth will be used in the calculations of the final teacher rating for the year.

Gary said, when speaking about the achievement part of the PAR system:

I am a junkie when it comes to scientific data. I have spreadsheets after spreadsheets. We do a pre and a post for each unit, and I keep track of test scores and questions they missed and areas of concern It helps me specifically with math; there is such a wide gamut of ability when it comes to math. . . . So that data helps me say, you need help with this, let me help you with this, or you are going to do this assignment because you need to work on this. It helps me with individualization and differentiation. It makes it more specific to that kid. My classes, I do different things in each class because each class is in a different place and have different things they need. That data really helps me.

Hillary provided the perspective of a teacher on the PAR committee and said the committee very carefully looked at the data component so it would be most helpful. She said that the administrators “were very honest with data portions, and the difficulty of reaching goals in every class was discussed. We determined here to being clearer on what type of assessments we as a staff should be putting in our growth data,” and this fine-tuning was another positive change.

Ian also talked about how the student data component was tied to growth. The teachers need to get 80% growth for a highly effective rating, and then input their data into PIVOT, and Ian has a discussion with them about where they fell. More experienced teachers help newer teachers with how achieve the student growth needed.

Gary said:

PERCEPTIONS OF TEACHER EVALUATION MODELS

Yes, my strategies have evolved as I am much more active in the data collection, a requirement of the PAR model, and the emphasis I place on that data for my instructional objectives. Additionally, because of my desire to track student data, I have chosen not to use the same lesson plans from year to year because of the data from the student's would be different.

Ian echoed that belief, "The data driven portion is fairly prescriptive in the fact that if you teach in a subject that has state or national testing then the data will be determined from the test scores." If the subject taught does not have standardized tests to use, then both teachers agreed they are allowed to create their own assessments for pre-tests and post-tests. Gary said, "The data achievement portion has continued to push me to have high expectations and goals. Evaluations drive better performance because it is measured." Similarly, Hillary believed, "The data achievement portion on PAR has been huge because the whole staff is trying to improve on their data and their ability to track that data." She elaborated:

I think under the PAR model we have all grown in our ability to track data and use it to help our instruction. Many times new teachers come in without the data component and under this model they get a clearer picture that expectations come when data goals are set and tried to be obtained on a yearly basis.

According to all three the last component of the data portion that has been a huge success factor was the belief that "growth is growth." This allowed the teachers the flexibility to be professional in their determination of how and where students can grow based on their level of knowledge. Also, the data achievement portion has evolved and been clarified well to serve the needs of the teaching staff. Thus, while Ian said he did not think student achievement had

PERCEPTIONS OF TEACHER EVALUATION MODELS

reached the level he desired, the fact that the data showed increasing growth had occurred was agreed upon by all three participants.

Lincoln High School: Themes that Emerged from Research Question 3

The third set of themes emerged as a result of data collected from participants' responses to the second research question: In what ways do teacher and administrator perceptions of the PAR evaluation model influence the evaluation process and educational outcomes? Two themes emerged: PAR was generally seen as a favorable change, so it was being well used by the teachers; and the manner of observations was not seen as uniformly positive.

Theme 1: PAR was liked, so followed

The first theme that emerged was that PAR was seen by most as a favorable change, so it was being well used by staff. All three participants corroborated the viewpoint that the staff liked the PAR model, especially if they had any experience using a different model, and all three specifically mentioned feeling that PAR empowered teachers much better than the previously-used RISE model, which they felt teachers saw as punitive. Hillary indicated that the PAR model has served JHS as "a fantastic model" with clear expectations and with "opportunities to change if the teacher wants to do so." The reason Ian gave for liking the model was that "the PAR model ties directly to what teachers are doing in their classroom. Also, there is not much pressure for teachers using this model, instead just an interest in getting better." Ian said he thought "it has a lot to do with our perception when [evaluators] go into a classroom," describing things that may once have seemed like a negative (e.g., noisy classrooms) that are now seen as a positive when evaluators' training kicks in and they evaluate using the rubric.

Ian has received no complaints verbally or in writing about the PAR model. Hillary stated similarly, "Staff appreciate the model. The process works through the entire whole

PERCEPTIONS OF TEACHER EVALUATION MODELS

structure of the PAR model. I have had no word from the staff about being unhappy and this has been great for the staff morale.” Hillary said generally teachers feel good about the model so when evaluators come into the room, they are okay with getting feedback; they want to improve. Specifically, Hillary said, “The model pushes me to do my best more so than others because I believe in the model.” Gary also indicated, “I like the PAR model, and I have grown because I find the feedback and reflection to be very valuable. The teacher gets out of the PAR model what they put into them.” Gary talked about being the type of person who liked the challenge he sees the rubric as setting up. His example involved using the rubric to make sure he had the data (could prove the student achievement) to back up being worthy of a good rating.

Part of what all three saw as contributing to the success of PAR was the buy-in from the staff. Prior to the implementation of the PAR model, a committee was formed to look at examples of PAR evaluation model and classroom observation documents that had been used in other districts, and they decided what was important and what they wanted to be seeing in the classrooms. Hillary said, “Having the teachers involved from the beginning was a game changer with buy-in and being able to hold firm on those expectations.” She added that “[t]eachers have bought into the PAR model because they think it is attainable.” All three indicated that PAR is the model where both teachers and administrators work together and professional growth has taken place through trust and openness in a collaborative process

One side note is that both of the teachers did indicate a small sense of a difference between newer and veteran staff. Specifically, Gary stated, “The model has helped me personally, but I sort of have the feeling that the veteran teachers still have evaluations in their mind as a check box item to be completed. As a newer teacher, I do not know any differently.” Hillary admitted that as veteran teacher she sometimes had to remember to look at the specifics

PERCEPTIONS OF TEACHER EVALUATION MODELS

of the rubric itself, citing the fact that she'd not looked back at it prior to the interview with me. She was reminded that "a few years into implementation we must keep it at the forefront of our vision and mission for the school."

Theme 2: Concern about timing

The second theme was that there was a perceived concern related to the time administrators have to do evaluations. Gary stated it this way:

Administrators only have so much time, and we only have so many prep periods or before/after school time to schedule the post conferences done. I can only imagine what their schedules look like trying to get every teacher's evaluation done throughout the school year.

And because this occurs, both teachers agreed that a difficulty for getting it all done resulted in sharing the load for administrators causing evaluations to be done by those who have never taught the subject matter. Hillary agreed, "I think finding time to give to give adequate feedback heading into the post conferences with the sizes of the staff while still running a building is always a challenge." All three agreed with the challenge of timing in all aspects for proper evaluations to take place. An example given was the difficulty having the time to reschedule observations and post conferences when something comes up for the teacher or the administrator.

The time factor is also compounded when there is teacher turnover and more time is needed with new teachers for proper implementation into the PAR model. Still, Hillary recognized that "the hope is to prevent the attitude of let's just get the evaluations finished and off the checklist and simply let the teachers do their thing in the classroom." Ian summed it up by saying, "Time to do it all, time to have discussion questions, and time to have productive post

PERCEPTIONS OF TEACHER EVALUATION MODELS

conference meetings. I think at times just the overwhelming nature of the spring semester becomes a challenge for both teachers and administrators with all aspects of evaluations.”

Relatedly, all three participants thought that differences in evaluators and issues with timing could have negative effects on evaluations. Basically, there was a feeling from both teachers that no matter how much time, effort, and detail go in to the training and coordination of the model, it felt like there was still a slight difference between evaluators and their personal expectations of teachers. Perceptions about certain teachers and how they were rated and marked on the rubric between different evaluators still continued to be a challenge using the PAR model. Hillary stated, “Sometimes it is amazing to me what administrators see and what they don’t see. Human judgment will always come under question and scrutiny.”

Lincoln High School: Summary

In this chapter, the findings from Lincoln High School (Case Study #3) were reported in order to answer the specific aspect of the overarching research question relevant to LHS: from the viewpoints of both teachers and administrators, does the PAR teacher evaluation model actually, in practice, cause teacher growth and improve classroom instruction? The findings of the three interviews were coded and themes were determined as they applied to the study’s three research questions. With regard to research question #1 pertaining to characteristics of the PAR model that the participants identified as most effective in promoting professional growth, four themes emerged. First changes to instruction occurred both in terms of better planning and preparation taking place prior to class and specific improvements made to classroom practices due to the use of PAR. Second, the quality and quantity of feedback improved under PAR. This was discussed regarding the post-conference feedback but also the assistance given by peers through the mentor teacher aspect of PAR. Third, under PAR, the participants felt that they were

PERCEPTIONS OF TEACHER EVALUATION MODELS

adequately trained initially and that they were re-trained appropriately. Finally, the fourth theme was that PAR provides common and clear language for shared expectations and this meant that everyone knew what was expected and how to best use the model to achieve growth.

With regard to research question #2 in which the participants were asked to identify the characteristics of the PAR model that were the most effective in promoting student achievement, two themes emerged. First, it was clear that using PAR caused the participants to change instructional strategies so that student learning became the focus of classroom teaching. Second, given the use of student data that was required under PAR, the participants were able to see, evaluate, and use student achievement data to further increase achievement.

Finally, with regard to research question #3, which explored the ways that participant perceptions of the PAR model influenced the evaluation process and educational outcomes, two themes emerged. First, because PAR was generally seen perceived as a favorable change, it was being well used by the teachers. Specifically, the participants noted that the process seen as collaborative and helpful and not threatening and led staff to want to improve. Finally, timing limitations (there are only so many evaluators with only so much time to evaluate all of the teachers) were a downside that was addressed. However, the recognition of this difficulty did not seem to detract from the overall positive feelings reported in the first theme.

Chapter seven follows and reports the result of my cross-case analysis of the major themes from Case Study #1 (Washington High School), Case Study #2 (Jefferson High School), and Case Study #3 (Lincoln High School) that were identified in chapters 4 through 6.

PERCEPTIONS OF TEACHER EVALUATION MODELS

CHAPTER SEVEN: CROSS-CASE ANALYSIS

In this chapter, I further analyzed the themes and patterns developed from my interviews, the data pertaining to the evaluation tools under each model, and the general observations made during the visits to all three high schools as written about in chapters 4, 5, and 6. The themes in chapter 7 are a result of cross-case analysis. Six themes that emerged from this cross-case analysis related to the three research questions concerning evaluation-based professional growth, evaluation-based student achievement, and the influence of participant perception on the evaluation process and outcomes. During this cross-case analysis, I looked for commonalities among all three models. Because my overarching purpose was not to compare and contrast the three models, I tried not to draw conclusions from differences in the models, but instead looked to similarities.

Before reporting the themes that emerged from my cross-case analysis, it is appropriate to present a cross-case comparison of the structure of the three evaluation models. Table 7 provides a direct comparison of key data.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Table 7

Comparison of key data in WHS, JHS, and LHS models

	Washington HS	Jefferson HS	Lincoln HS
Percentage of summative evaluation allocated to classroom observation	75	85	60
Percentage of summative evaluation allocated to student growth	20	0	40
State letter grade for school included?	Yes	No	Yes
Number of classroom evaluations	5	3	4
Announced/Unannounced	2/3	2/1	2/2
Post-evaluation conference required?	Yes	Yes	Yes
Prior practice	Yearly goal setting	An evaluation every other year	An evaluation every year

The three different models had critical similarities. First, all three models employed at least three evaluations, all of which were full-period evaluations. Further, JHS had just changed from four evaluations to three this year because, after trying the process with four evaluations, it determined that three would serve the purpose of the evaluation system and would be easier to facilitate. The number of evaluations per year was a dramatic increase from pre-RISE/TAP/PAR days, which at most involved a single yearly observation. Having abundant evidence from numerous evaluations in a school year indicates a thoughtful approach to teacher evaluation—one that engages teachers in reflection and self-assessment (Danielson, 2011).

All three models required a post-conference after evaluations. This translates to teachers getting direct, timely feedback from evaluators. A clear indicator of the strength of each model was the fact that well over half (60-85%) of the summative evaluation was centered on the classroom observation piece, and while there were differences in component parts outside of

PERCEPTIONS OF TEACHER EVALUATION MODELS

classroom observations, a full 85-100% of the total was the combination of classroom observation and student growth for each school.

There were three themes that pertained to research question #1: (a) evaluation models facilitate professional growth when they include training; (b) evaluation models facilitate professional growth when they cause positive change to instruction; and (c) evaluation models facilitate professional growth when they include feedback. There were two themes that pertained to research question #2: (a) evaluation models promote an increase in student achievement when they support teaching to be more student focused; and (b) evaluation models promote an increase in student achievement when they incorporate student data. Finally, there was a single theme pertaining to research question #3: evaluation models that are perceived positively are more likely to be used by teachers, and thus, have a greater influence on processes and outcomes.

Themes Pertaining to Professional Growth

Theme #1: Evaluation models facilitate professional growth when they include training

The first cross-case theme that emerged was that an evaluation model facilitates professional growth when it includes training. Each of the schools included training for its evaluation model at the beginning of each school year. This training was more intensive the first year of implementation, but it continued at each school at the start of every year. Under all three models, the initial training, undertaken by the entire staff, addressed all of the model's rubric components individually. Then, in following school years, there was both the same sort of all-encompassing training for new (or new to the district) teachers and retraining and working on question areas for all teachers. In addition to this formal training, the teachers at all three schools also talked about the value of the informal training that happened under their respective models.

PERCEPTIONS OF TEACHER EVALUATION MODELS

The use of all models led to some component of sharing (e.g., techniques, methods, practices) between staff members.

At WHS under the RISE model, training allowed everyone to understand the rubric and retraining was used as a refresher to keep everyone up to date each year. The value in consistent language used in evaluations every year meant that both administrator and teachers felt it relatively easy to engage in training and use the system to continue to improve. Weekly PLCs provided the more formalized training ground for not only the rubric but also the SLOs and TLOs, and informal training routinely happened as needed between department heads and teachers and even between teachers and teachers. None of the participants expressed any concern with the amount of training, and Adam stated the consensus well when he said the training was just what was needed for teachers to use RISE to improve.

At JHS, under the TAP model, the theme pertaining to feedback was similar. Original training and continued review/training were key aspects of the model that, according to Freda, “served to give the feeling for those trained that [they] were all in [it] together.” TAP included weekly cluster meetings and those meetings were used for training purposes early each year. The participants all agreed that breaking down the rubric and working on determining best practices as related to the rubric helped bring about better teaching. The informal training at JHS came via the established master and mentor teachers. Not only did those individuals specifically assist with training (and evaluating teachers in the case of the master teachers), it was clear that they were also valued resources for general conversations between teachers related to classroom practices. These collaboration opportunities were seen as “critical” to professional growth.

Lastly, the PAR model as used at LHS also uses training and retraining to keep teachers pursuing personal improvement. There seemed to be a consensus that the participants valued

PERCEPTIONS OF TEACHER EVALUATION MODELS

training as a necessary aspect of using an evaluation system so that the enthusiasm to master the rubric would be carried over from year to year. Particularly important at LHS was the feeling that training showed participants that the PAR expectations were attainable, and this belief was very important when it came to facilitating growth.

All nine participants at the three schools agreed that training was a component of the three different evaluation models that facilitated professional growth. Research shows the need for training for evaluators and for teachers to understand what constitutes great teaching (Danielson, 2011). This first cross-case theme clearly showed an understanding that trained evaluators better understood what to look for in evaluations and trained teachers better understood what was expected and what to strive toward when being evaluated.

Theme #2: Evaluation models facilitate professional growth when they promote positive change to instruction

The second cross-case theme revealed that the participants at all three schools felt like changes were made to instruction under the respective models that contributed to professional growth. A few general principles emerged: the model rubrics caused evaluators to change to a much clearer, more focused, and more purposeful means to observe instruction; the models gave teachers the means and the motivation to prepare for class better than they had previously; and the rubrics pinpointed good teaching techniques, and teachers used them to change what they were doing in the classroom for the better.

The three participants at WHS spoke at length about how using RISE caused them to change what they were doing on a nearly daily basis. They believed that having the teacher and the administrator on the same page made everyone more confident in knowing what needed to be or could be made better by being changed. That this coupled with seeing the positive results of

PERCEPTIONS OF TEACHER EVALUATION MODELS

changes to instruction were powerful motivators for working to improve classroom instruction.

Both Adam and Betty could easily bring to mind specific changes they had made in their teaching that they attributed directly to RISE, and Carter recalled with pleasure specific instances of pointing teachers to rubric language that was incorporated into better instructional practices.

Similarly, at JHS Danielle and Evan refined their teaching methods, and Freda improved her evaluation practices, which she felt contributed directly to improved instruction by her teachers. As with the other schools, at JHS the participants believed their model resulted in greater intentionality in the classroom and that this started with more purposeful planning. Even teachers who are already doing well could use TAP to validate their current practices, reinforce what is working, and look for nuanced ways to improve.

Finally, the LHS participants articulated this theme pertaining to improved instruction in a similar manner. The only slight variation under PAR was that Gary, Hillary, and Ian were even more vocal about the improvement to instruction involving the collaborative use of the model. At LHS, there were certain improvements that were brought about because of the model rubric, such as bell ringer and exit ticket strategies, that were being incorporated nearly school wide. This, of course, was in addition to individual teachers deciding to change things as appropriate in individual classrooms.

It is significant that all participants—at three schools using three different evaluation models—concluded that professional growth by means of improved classroom instruction took place for teachers and administrators. This indicates that the overall effect of evaluations being done annually with specific rubrics post-P.L. 90 has been to positively change instructional methodology. This is in keeping with Danielson's (2011) conclusion that this sort of professional development is a critical aspect of evaluations. As she said,

PERCEPTIONS OF TEACHER EVALUATION MODELS

A commitment to professional learning is important, not because teaching is of poor quality and must be ‘fixed,’ but rather because teaching is so hard that we can always improve it. No matter how good a lesson is, we can always make it better (p. 37).

Theme #3: Evaluation models facilitate professional growth when they include feedback

A third theme that emerged from the cross-case analysis with the interviews was that the time teachers spent with evaluators, most often in post-conferences, for feedback and reflection provided growth for staff members. A post-conference was scheduled between administrator and teacher after every observation under all three models. Additionally, the models used a combination of other types of feedback, including pre-conferences and working with mentor teachers. Regardless of exactly how it was done, the staff taking time to sit down and have professional conversations was perceived as contributing to growth. This validates research showing that feedback and dialogue given in a manner that helps guide improvement and best practices in the classroom is one of the most critical components of an effective evaluation model (Bambrick-Santoyo, 2012; Danielson, 2012; Darling-Hammond, 2015; Marzano, 2015).

An overview from the three participants at WHS, which used RISE, indicates that when the post-conference meetings were held, they knew that, as best summarized by Adam, “coaching was going to be the main theme. The observation would be critiqued and then open-ended discovery questions would be asked to enhance dialogue.” In addition, teachers were given things to think about in post-conferences that they later implemented to benefit student learning. Finally, post-conferences were said to be the most critical part of the evaluation process at WHS, allowing each participant the opportunity for collaborative, reflective conversations with each other that lead to a heightened awareness of other people’s perspectives. Growth occurred from the feedback and dialogue between the evaluator and teacher which

PERCEPTIONS OF TEACHER EVALUATION MODELS

provided a common vision and set of expectations for the classroom. The RISE model expectations for these conferences was perhaps the thing that helped teachers grow the most.

Comments from the participants at the school using TAP followed the same thematic pattern. All three participants agreed about the importance of feedback and dialogue in pre- and post-conferences. In these sessions at JHS, teachers got to do most of the talking and administrators ask needed directional questions in order to see areas for refinement and reinforcement. The conferences allow for discussion centered on indicators listed in the rubric. An example was provided where changed teaching practices in the classroom were completed based on feedback provided by the evaluator. The post-conferences offered suggestions or nuggets of knowledge from the evaluators who have more experience and knowledge in the educational arena. JHS also uses master and mentor teachers as a formal part of the TAP system, and feedback from those teachers was seen as very valuable. The idea of working and collaborating together in conferences to discuss needs and components of the evaluations has brought alignment with teachers and evaluators. Growth in the final summary ratings have occurred as a result of the feedback provided during the conferences.

The participants at LHS, which used PAR, provided similar comments about feedback and dialogue. They reported that having the requirement to sit down face-to-face for professional conversations as a part of every observation provided growth opportunities for the teaching and administrative staff. The conversations and feedback pushed and challenged them to get better in classroom instruction, and caused teachers to “want that challenge; want that feedback from [their] evaluator[s],” according to Hillary. They also saw the use of mentor teachers as a “huge selling point” of the model, as characterized by Gary, which they saw as a way for tough conversations to be more palatable to teachers and, thus, more likely to lead to

PERCEPTIONS OF TEACHER EVALUATION MODELS

actual growth. Furthermore, because of the required post-conference conversations, the evaluation feedback was not finalized until after the post-conference allowing the opportunity for the teachers and administrators to understand the specifics of the model more clearly when they went over it together. This conferencing time allowed for asking questions, seeking clarity and wisdom, and explaining their thinking, if necessary. At LHS, there was a feeling created by the conferences that the evaluators and evaluates were in the process together and all involved saw that as being tied to growth.

According to Danielson (2011), the evidence is well established that a purposeful approach to evaluation (such as being used under all three models) requires teachers to be engaged in feedback, reflection, and self-assessment and yields benefits that result in teacher growth rather than merely meeting the goal of ensuring they are of good quality. This is the result of agreed-upon standards providing the chance for participants to be engaged in professional conversations. The data from the participants at all three schools/models were entirely consistent with the research that according to Goe (2013) consistently shows the importance of timely feedback in the evaluation process. Goe cited a particular study that showed a connection between principal feedback and instructional improvement (2013), and the results of the three Case Studies certainly mirror that finding.

Themes Pertaining to Student Achievement

Theme #1: Evaluation models promote an increase in student achievement when they support teaching to be more student focused

During my cross-case evaluation process, I saw that during the interviews the participants discussed how the models led each school to become much more concerned than they had been both with what the students were doing (how they were learning) and with what the students

PERCEPTIONS OF TEACHER EVALUATION MODELS

were learning (how they were performing). There can be no question that the overall best practice for effective instructional strategies now centers on student engagement—a change from mostly lecture to more discussion-based, hands-on learning (Darling-Hammond, Barron, Pearson, Schoenfeld, Stage, Zimmerman, & Tilson, 2015; Laurillard, 2013). All of the participants corroborated this thinking and, in fact, indicated that all three models, specifically their indicators/components, do align with the current thinking on student engagement. Some examples of the changes resulting from the evaluations' rubric components include: (a) changing instructional strategies with experienced teachers' methods evolving; (b) getting students to do more grouping and having them be active participants; and (c) increasing student-to-student interactions and students asking questions to increase learning opportunities. These examples helped shape and mold the administrator and teacher feelings that the models aligned with good lessons and that changes were made from the evaluation protocols to improve teaching practices in a manner that was student focused. In all three schools, the participants agreed that the models aligned with good lessons and well-run classrooms and that every indicator in the models lined up best teaching practices and what is best for students.

At WHS with the RISE model, the evaluations led them as a school to become much more concerned with what students are doing and how they are learning and less concerned with what was easier for the educators. A central competency for instructional strategies now focuses on student engagement and checking for understanding. Teachers indicated that they were much more student centered and that they have changed from mostly lecturing, sit-and-get, to a more discussion style, hands-on learning atmosphere. Changing the way teaching has been performed has shaped the way activities are created and formative and summative assessments are written, all to the benefit of student learning in the classroom.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Participants who used TAP model similarly believed that the model facilitated greater student investment and involvement in learning that led to an increase in student achievement at JHS. The model has resulted in a more student-centered classroom. Students used to be sitting quietly at their desks doing work individually or listening to teachers lecturing, compared to now-norm of student-to-student interactions and questioning. More active learning is occurring because of the model expectations. In addition, students were seen as clearly benefitting from changed instructional practices, which have aided the students who now are challenged to reach deeper and engage in higher order thinking. For instance, grouping and getting students to actively work is a strategy that was incorporated because it is written as a competency in the TAP model. For all, reaching that competency of the model rubric has been the goal, and increased student achievement has resulted.

Finally, participants at LHS using the PAR model presented a similar theme involving the purposeful fostering of student achievement. Once again, it was very clear that all three participants had, in fact, used the rubric to be more focused on using tactics that directly affecting student achievement. For example, the PAR rubric has helped them try to increase engagement, because they are purposeful about having all students engaged. That change came directly from the PAR model component which speaks about the development of skills by using hands-on activities with the students. As a common theme, the PAR school also indicated that checks for student understanding are now prevalent in the classrooms because of the categories on the evaluation rubric. From an administrative perspective, there was confirmation that there is less whole-group-lecture style of teaching happening and much more focus on guided instruction and best practices for student growth.

PERCEPTIONS OF TEACHER EVALUATION MODELS

All participants believed that the changes that the various models required be made in order to increase student engagement were actually being made and that doing so increased student achievement. Danielson (2011) confirmed that evaluation systems allow for a common language to describe desired practices (like increasing student engagement in learning), and this practice causes feedback to be more likely to be implemented, which brings about positive change to student achievement.

Theme #2: Evaluation models promote an increase in student achievement when they incorporate student data

Initially, it appeared that because only two schools used data in their models (WHS/RISE and LHS/PAR) and one school (JHS/TAP) did not, the use of data would not emerge as a theme. However, upon further reflection, I found that the two opposite uses of data advanced the same principle: it was a positive to increasing student achievement to utilize student data, or saying the same thing from the opposite perspective, it was a negative to not use it.

The RISE model required the use of student data, and the three WHS participants all felt that when discussions about using data occur, great student achievement is possible and more productive. All three participants clearly believed WHS student achievement was improved by the use of data in the accountability/evaluation process. In fact, the area where RISE has made an enormous difference at WHS is in the requirement to have SLOs and the expectation to track and report the data. This allows all involved to have goals and expectations that are clear and measurable, and it allows for administrators to work with teachers so that their goals are appropriate and attainable while providing the needed push for improvement. Given the data portion requirement, WHS has aimed to ensure that teachers perform at their best to enhance student learning. This led to a greater sense of accountability under the RISE model.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Under the PAR model, the requirement to use of student achievement data was also seen as fostering student achievement. The overall belief was that the data component was one of the characteristics of the PAR model that is beneficial because it requires data be used to drive instruction and be shown in the evaluation process. Teachers are expected to keep track of the data and report it at the appropriate time. Data growth is used in the calculations of the final teacher rating for the year. Things are done differently in each class because each class has different data points, and experienced teachers help newer teachers with how to achieve the student growth needed. Because of this, lesson plans are changed from year to year because the data from the students would be different, and this effect of data use is viewed as a clear contributor to achievement growth.

The thinking under the PAR model is that all of the staff has grown in the ability to track data and use it to help instruction. Many times new teachers come in without the data component, and under this model they get a clearer picture that expectations come when data goals are set and tried to be obtained on a yearly basis. This allowed the teachers the flexibility to be professional in their determination of how and where students can grow based on each student's level of knowledge. Additionally, the data achievement portion has evolved and been clarified well to serve the needs of the teaching staff.

Alternatively, the TAP model used at JHS does not involve the use of student data. JHS does not require teachers to track, use, or report data on student achievement. While data is apparently looked at occasionally in a few circumstances (e.g., department meetings), teachers were not responsible to provide and administrators did not ask for student data. One teacher, Danielle, said that for the most part, evaluation discussions around data have been “more anecdotal and individualistic in nature and nothing has to be turned in as a teacher to be part of a

PERCEPTIONS OF TEACHER EVALUATION MODELS

final percentage of your yearly review.” This meant that when participants were asked about student achievement, they could not answer in specific terms. Instead, they were forced to rely on generalities such as an increasing graduation rate or a better state letter grade. This resulted in, as Evan stated, this issue being “tough . . . to address” because there was no data that “show[ed] that student achievement has improved,” and thus there could only be an “assumption” regarding student achievement being tied to the use of the TAP model.

The three JHS interviews supported the finding that administrators and teachers would know better how the evaluation model was or was not specifically impacting student achievement if the TAP model would use data in some fashion. And, this seemed entirely consistent with what was being said at WHS/RISE and LHS/PAR and with research generally, given that it is understood that use of student achievement data that is tailored to a specific school’s practices (teachers have a dialogue with evaluators as to the assessment tool to use) should make up a portion of the teacher evaluation model (Marshall, 2012).

Theme pertaining to the influence of perceptions on process and outcomes

Evaluation models that are perceived positively are more likely to be used by teachers, and thus, have a greater influence on processes and outcomes

There emerged an overall sense from all of the interviews that while teachers generally seem to have bought into the use of each of the models, new teachers mind it less because it is all they have ever known and experienced teachers find it to be a bit more of an inconvenience. However, despite this perception concerning experienced teachers, the three experienced teachers all were individually forthcoming that they are better teachers than they were before the evaluation models set the annual expectations. This led to an overall acceptance with both the administrators and teachers realizing the models are here to stay, and each group has gotten more

PERCEPTIONS OF TEACHER EVALUATION MODELS

comfortable with the procedures. In sum, the staffs appreciated the models and will get benefit from them if they are willing to invest in the process.

With the RISE model at WHS, the participants felt that the more positive the perceptions held about RISE the more likely the teacher or administrator was to subscribe to its principles and use it as a basis for improvement professionally and for student achievement. They believed that some experienced teachers still see being evaluated as just another thing to accomplish on the yearly task list and are a bit more hesitant or resistant to feedback on practices that have been occurring for many years. But, the participants conveyed the sense was that there were not many of those teachers, and for those teachers who want feedback, who want to grow, and who want to improve, the RISE system is excellent. Collectively, Adam, Betty, and Carter felt that overwhelmingly positive feelings about RISE has led to its process being accepted and implemented and desired outcomes being obtained.

In JHS's TAP model, the similar perception was present: there is a certain degree of buy-in to the evaluation model that influences teachers' desires to use it for improvement. The positive feelings regarding the model stemmed from the belief that the model completely aligned with good lessons and a well-run classroom. In fact, the participants felt that the model's indicators are lined up with best teaching practices and what is best for getting students to grow. In practice, the staff believes that the rubric domains and indicators are aligned with one another, meaning that if a teacher excels in one part of the rubric, he or she will most likely excel in other parts. An example provided by Evan indicated that a teacher's knowledge of students could impact lesson structure and pacing, which then impacts measuring student mastery. There seemed to be a combined perception that more of the JHS staff see TAP as a "necessary evil" than was present at the two other schools, but the three participants in the end agreed that

PERCEPTIONS OF TEACHER EVALUATION MODELS

teachers have accepted it and are using it to be more reflective in their craft, making them assess decisions and actions in the classroom more deliberately.

The perceptions of the LHS participants, who all used the PAR model, were that the evaluation model is more about getting better at what they were doing rather than performing in a specifically expected manner. The LHS theme centered on the perception that the PAR model was more concerned with what was actually being taught in the classroom than micromanaging classroom management techniques. The teachers felt that generally evaluators without the rubric just looked at the outside, to see how things seemed to be going, but under PAR evaluators were able to look for more things and use the data to see that growth is actually happening. The consensus was that the model had them trying to focus on the growth aspect instead of a teacher having to be perfect in every category right from the start. LHS teachers have bought into this PAR model because they think it is attainable. Buy-in also occurred given the feeling that the evaluations were more about meeting the needs individual teachers in the classroom rather than looking at the evaluation as a checklist to be mastered. The PAR model has helped LHS teachers because they have more interest in getting better under this model.

In all three schools and under all three models, the belief that the evaluation system was for the individual and common good of the teacher and the school was critical to the models' success. In varying terms, all described the process as one that was not out to get teachers and one in which evaluators and evaluatees worked together. This aligned with Bambrick-Santoyo's (2012) conclusion that "[w]hen teachers see the concrete steps they must take to improve their practice, and when they can continually practice skills connected to those steps, transformational success come within reach" (p. 30).

PERCEPTIONS OF TEACHER EVALUATION MODELS

Summary

This chapter outlined the results of the cross-case analysis of the three individual cases. Six overall themes emerged. There were three themes that pertained to research question #1: (a) evaluation models facilitate professional growth when they include training; (b) evaluation models facilitate professional growth when they cause positive change to instruction; and (c) evaluation models facilitate professional growth when they include feedback. There were two themes that pertained to research question #2: (a) evaluation models promote an increase in student achievement when they cause teaching to be more student focused; and (b) evaluation models promote an increase in student achievement when they incorporate student data. There was a single theme pertaining to research question #3: evaluation models that are perceived positively are more likely to be used by teachers, and thus, have a greater influence on processes and outcomes.

Chapter 8 follows, and in it these cross-case findings will be measured against the six utility evaluation standards developed by The Joint Committee on Standards for Educational Evaluation (2009) as the culmination of my research. The connection of my findings with the Utility Standards will answer my research questions: What characteristics of teacher evaluation models do teachers and administrators perceive to be the most effective in promoting professional growth? What characteristics of the teacher evaluation models to teachers and administrators perceive to be the most effective in promoting student achievement? In what ways do teacher and administrator perceptions of the evaluation models influence the evaluation process and educational outcomes.

PERCEPTIONS OF TEACHER EVALUATION MODELS

CHAPTER EIGHT: CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

This chapter summarizes the study of Indiana teacher and administrator perceptions regarding three different teacher evaluation models, RISE, TAP, and PAR, and examines conclusions from the data presented in chapters four through seven. A summary of the study includes an overview of the problem, the research questions, a review of the methodology, and a discussion of major findings in the context of the literature and previous studies discussed in chapter 2. Then, unexpected findings in this study are discussed, followed by conclusions. As a part of the conclusions, implications for practice and recommendations for further research are highlighted. This chapter ends with concluding remarks that connect the findings to the greater body of educational research.

Summary of the Study

The purpose of the study was to analyze the perceptions of teachers and administrators from three different high schools regarding their three different teacher evaluation systems—RISE, TAP, and PAR—with particular emphasis on the characteristics of the models that promote professional growth and student achievement. The three high schools selected for this study began implementation of the new teacher evaluation mandates in 2012-2013; thus, my interviews in the fall of 2017 meant the schools had five years with the new expectations and requirements in place at the time data was collected.

Overview of the problem

The problem is that schools are complex, dynamic systems that influence students' academic success, and the evaluation models they employ need to be constantly scrutinized from multiple perspectives to determine if teachers and administrators perceive the evaluations to enhance their ability to promote professional growth and improve instruction. Teacher

PERCEPTIONS OF TEACHER EVALUATION MODELS

evaluation reform has clearly happened over the last five years, and what remained to be answered was whether this course of action has changed the teacher evaluation process for the better from the teacher and administrator perspective.

I compared teachers' and administrators' perceptions of the extent to which these new evaluation models increased support for teacher effectiveness, raised student achievement accountability, and brought about desired professional growth by Indiana teachers and administrators. The information generated from the results of this study contributes to the data available to Indiana school districts making decisions about their current evaluation practices. Furthermore, the comparison of this cross-case analysis could be useful in understanding specific impacts of the evaluation models on teachers and administrators.

Research questions

In this chapter, the themes that emerged from my cross-case analysis were compared against the six utility evaluation standards as the culmination of my research to answer the overarching question for my research: Whether the three teacher evaluation models—RISE, TAP, or PAR—cause teacher growth and improve student achievement from the viewpoints of both teachers and administrators. The following specific research questions were used in this study:

1. What characteristics of teacher evaluation models do teachers and administrators perceive to be the most effective in promoting professional growth?
2. What characteristics of the teacher evaluation models do teachers and administrators perceive to be the most effective in promoting student achievement?
3. In what ways do teacher and administrator perceptions of the evaluation models influence the evaluation process and educational outcomes?

PERCEPTIONS OF TEACHER EVALUATION MODELS

Review of the methodology

This qualitative study was based on nine participant interviews that were completed in the Fall of 2017. I interviewed an administrator and two teachers from Washington High School that used the RISE teacher evaluation model; an administrator and two teachers from Jefferson High School that used the TAP evaluation model; and an administrator and two teachers from Lincoln High School that used the PAR evaluation model. The three schools were selected because they each used one of the three different evaluation models that I was seeking to study in the course of my research. An interview protocol with vetted and approved questions was created (Appendix A), and using those questions, I interviewed the nine participants in person in their schools.

During a three-month period, data was collected in the form of the transcribed interviews, my personal observations, and the documents setting out the specifics of the evaluation model used in each district. I analyzed the data individually in each of the three case studies using open coding procedures. The coding allowed me to identify major themes in the data. Then I engaged in a second cycle of coding during which I pared down my original, general determination of themes into themes that specifically pertained to my three research questions. This process left me with eight themes in Case Study 1, seven themes in Case Study 2, and eight themes in Case Study 3. Next, I undertook a cross-case analysis. I compared the data from three cases to determine the final themes that emerged across the case studies in order to answer the three research questions in this study. There were six themes that emerged from the cross-case analysis.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Major findings

The first three themes pertained to research question #1: (a) evaluation models facilitate professional growth when they include training; (b) evaluation models facilitate professional growth when they cause positive change to instruction; and (c) evaluation models facilitate professional growth when they include feedback. There were a fourth and fifth theme that pertained to research question #2: (d) evaluation models promote an increase in student achievement when they cause teaching to be more student focused; and (e) evaluation models promote an increase in student achievement when they incorporate student data. Finally, the sixth theme pertained to research question #3: (f) evaluation models that are perceived positively are more likely to be used by teachers, and thus, have a greater influence on processes and outcomes.

Findings related to the literature

There are six assumptions that guided The Joint Committee on Standards for Educational Evaluation (2009) in organizing the 27 Personnel Evaluation Standards into categories based on four essential attributes. The second of those attributes is utility. I focused on the utility category because it was most directly applicable to my research questions because the Utility Standards “view personnel evaluation as an integral part of an institution’s ongoing effort to . . . provide [staff members] with timely and relevant evaluative feedback and encourage and guide them to deliver high quality service” (p. 6).

There are six standards. Utility Standard 1 requires the evaluation be constructive, in that it reflects the institution’s goals and missions and builds evaluatees’ “professional self-knowledge [and] increase[s] their enthusiasm and their efficacy as practitioners.” Utility Standard 2 requires that the evaluators consult, at the start of the evaluation process, with the

PERCEPTIONS OF TEACHER EVALUATION MODELS

intended user groups to determine the intended uses of the evaluation results and findings so that evaluations will be guided by their intended uses. Utility Standard 3 requires that evaluations being completed by credible evaluators who have the title, training, and/or certification to evaluate and the ability to understand the expectations of the position being evaluated. Utility Standard 4 requires evaluations to have clear performance expectations in words that are easily understood by the evaluator and evaluate, including job expectations and a rating scale or rubric that are clearly defined and explicitly aligned. Utility Standard 5 requires evaluations provide clear and timely results via written reports that accurately measure evaluatees' job performance, so professional development and personal growth can happen effectively. Lastly, Utility Standard 6 requires evaluations be used to improve staff performance as a whole so that the evaluation process leads to individualized and full staff professional development plans and programs. The six Utility Standards formed the conceptual framework of this study in that they guided my development of the three specific research questions under which my cross-case final themes were organized.

Before comparing the cross-case themes with the Utility Standards, I compared the structures of the three evaluation models—RISE, TAP, and PAR. This review revealed that the three models were in alignment with Utility Standard 4. All of the models had clear performance expectations (in words and percentages) that those using the model could easily understand and implement. This alignment was further supported when, after studying the teacher evaluation models in the documents provided to me by the school districts, it was clear that all three models had job expectations listed in the performance indicators as well as rating scales/rubrics that were clearly defined and explicitly aligned with the expected standards. This aligns with research that shows that educators are more likely to determine what makes teachers effective in the classroom

PERCEPTIONS OF TEACHER EVALUATION MODELS

when talented educational experts create comprehensive rubrics concerning what teaching should look like (Bambrick-Santoyo, 2012). All three evaluation models answered the charge made by researchers such as Danielson (2011) that teacher evaluation systems were not yet accurately measuring teacher quality and needed to do so. According to Danielson, the two goals for teacher evaluation are professional development and quality assurance. Danielson's framework seeks to provide detailed and objective criteria to describe teaching behaviors in evaluations. The intent is to reduce the subjectivity that is created by different experiences and perceptions among administrative evaluators. Finally, all of the evaluation models used in the three Cases were based on a growth mindset, something that the research uniformly finds necessary (Almy, 2011; Darling-Hammond, 2013; Marzano, 2015).

Research Question 1: What characteristics of teacher evaluation models do teachers and administrators perceive to be the most effective in promoting professional growth?

The first cross-case theme, related to research question 1, was that evaluation models facilitate professional growth when they include training. This theme aligns with Utility Standards 2 and 3. Utility Standard 2 involves evaluators understanding and communicating with evaluatees the intended uses of evaluation results, and Utility Standard 3 states that personnel evaluations should be completed by credible, well-trained evaluators who have the title and the training or certification to evaluate and to understand the expectations of the positions being evaluated.

In all three schools, training occurred for evaluators prior to implementation of the models and at the beginning of each school year. Under all three models, the training for both teachers and administrators occurred by addressing the model's rubric components. In addition to this formal training and re-training, informal training that led to greater sharing of knowledge

PERCEPTIONS OF TEACHER EVALUATION MODELS

also took place. Related to Utility Standard 2, it was clear that all involved understood exactly how the evaluation results would be used. This removal of any ambiguity and clear alignment of expectations, results, and outcomes was a specific purpose behind the legislated adoption of a uniform model for evaluation (Whiteman et al., 2011), and it was an aspect of all three evaluation models.

Additionally, all of the administrators addressed their beliefs that training on and use of the evaluation rubric gave them better credibility and ability to understand what to look for to assess the nuances of each teacher's craft, thus, aligning unambiguously with Utility Standard 3. Further, this theme pertaining to the value of training is also addressed in the Danielson framework, where it outlines the need for training of evaluators and a clearer understanding by teachers and administrators about what is great teaching (Danielson, 2011; Danielson & McGreal, 2000). Training needs to take place to make sure that the evaluations were authentic and that all involved were on the same page. Well-trained evaluators are vital to ensuring that well-designed evaluation instruments are implemented with fidelity (Dillon, 2013). Thus, both the research and the expectations from Utility Standards 2 and 3 that are fulfilled when training for teachers and administrators is a part of an evaluation model validate the conclusions stated in Theme 1 that models that include training promote professional growth.

The second cross-case theme centered on the conclusion that the studied evaluation models facilitated professional growth by causing positive change to instruction. This theme aligned with Utility Standards 1 and 6. Utility Standard 1 carries with it the expectation that teachers' effectiveness in the classroom is bettered, and Utility Standard 6 involves individualized performance improvements. Research makes it clear that no matter how many rubric categories teachers are proficient in, if there is no true method in place to seek to improve

PERCEPTIONS OF TEACHER EVALUATION MODELS

their competence in the classroom, then it only measures and it does not change the teachers (Almy, 2011; Danielson, 2012; Darling-Hammond, 2015; Marzano, 2015; Nolan & Hoover, 2011; Peterson, 2000). Related to the first standard, the second theme showed that all three evaluation models resulted in increased efficacy of practitioners which led to positive changes to instruction. Similarly, related to the sixth standard, all nine participants reported specific changes made to instructional methods as a result of the evaluation models. These were something more than the school-wide training or suggestions; these were the individualized performance improvements that Utility Standard 6 requires.

In the Sartain et al. (2011) study, nearly all teachers felt that their practice had improved due to the identification of the conferencing process as a critical aspect of the self-knowledge. Teachers reported improvement in planning, classroom management, using assessment during instruction, differentiated instruction, and student-focused learning. Here, many of these key strategies were discussed by the participants and could then be arranged in a holistic view of positive changes to instructional practices. Thus both the research and the expectations from Utility Standards 1 and 6 that are fulfilled when an evaluation model causes teachers to change what they are doing in the classroom validate the conclusion iterated in Theme 2 that models that promote changes to instruction promote professional growth.

The third cross-case theme was that evaluation models facilitate professional growth when they include feedback. This theme lines up with Utility Standard 5, which requires evaluation models to provide clear and timely results via reports that accurately measure evaluatee's job performance, and Utility Standard 6, which includes the giving of credible feedback from evaluators. As this third theme made clear, all three models used timely and meaningful feedback from evaluators based on a rubric that was used to accurately measure job

PERCEPTIONS OF TEACHER EVALUATION MODELS

performance. Each of the models consisted of forms that evaluators and evaluatees were familiar with and understood to be the means for tracking performance. Expressly stated in the rationale for Utility Standard 5 is that “[w]hen results are clear, timely, accurate, and germane to the purpose of the evaluation, follow-up actions such as faculty development . . . follow naturally and effectively” (JCSEE, 2009, p. 90). Additionally, Utility Standard 6 unambiguously states in its explanation that “[c]redible feedback from trained evaluators who base their judgment on the criteria of the evaluation system is crucial” (p. 94). Cross-case theme 3 shows these standards were being met.

This further aligns with the research, which shows that the use of feedback allows the evaluation process to improve the quality of teaching rather than serving only as an accountability mechanism (Danielson, 1996; Danielson & McGreal, 2000; Goe, 2013; Natriello, 1990). Good evaluation models “are designed as coaching models for teacher development, using skillful coaching and working with teachers on specific concrete actions that improve results” (Bambrick-Santoyo, 2012, p. 28). In essence, according to Bambrick-Santoyo (2012), regular feedback is essential, and the necessary loop of feedback, correction, and improvement cannot happen in evaluation models where feedback and reflection are not made high priorities through the building of trust and collegiality. All of the participants reported their models helping build trust and deeper relationships and contributing to the feedback loop. Additionally, studies have shown that peer assistance through formal or informal mentoring arrangements yields considerable teacher growth (Goldstein, 2008; Keller, 2006; Johnson & Fiarman, 2012; Munson, 1998; Papay & Johnson, 2012). The participants in this study made it clear that this reported value of peer assistance was true in their circumstances. Thus, both the research and the expectations from Utility Standard 5 that are fulfilled when an evaluation model requires

PERCEPTIONS OF TEACHER EVALUATION MODELS

evaluators and evaluatees to meet so that the reported results of the evaluation are understood validate the conclusion presented in Theme 3 that models that include feedback promote professional growth.

Research Question 2: What characteristics of teacher evaluation models do teachers and administrators perceive to be the most effective in promoting student achievement?

Pertaining to the second research question, the fourth cross-case theme was that the models promoted teaching to become more student-focused. Utility Standards 1 and 6 are relevant to this theme for reasons similar to those outlined with regard to theme 2. As discussed, Utility Standards 1 and 6 both include a component related to empowering teachers to make individualized changes to use best practices in their classrooms. As all participants reported that the evaluation rubrics included criteria related to student engagement, it was apparent that the expectation that teachers follow the rubric meant student engagement would increase if teachers were performing well. Importantly, the rationale for Utility Standard 1 says that strong evaluation systems accomplish characteristics such as “reinforcing strong professional practices” and “[e]ducators and institutions that achieve such beneficial characteristics are better able to provide high quality services to students and to maintain their effectiveness . . .” (JCSEE, 2009, p. 70).

The research stresses the importance of establishing professional practice rubrics and comprehensive teaching frameworks that define effective instruction to look more positively at teaching and learning (Frontier & Mielke, 2016; Pieczura, 2012; Schachter, 2012; Zatynski, 2012). This correlation between instructional strategies and student achievement is a focus of Marzano’s framework (Marzano, 2015). That framework includes specific elements that are scored in order to critique development and that establish a knowledge base that allows teachers

PERCEPTIONS OF TEACHER EVALUATION MODELS

and administrators to develop expertise. This was precisely what all of the participants reported their evaluation systems as doing. Thus it seems apparent that the expectations from Utility Standards 1 and 6 that are fulfilled when an evaluation model results in improved student-focused teaching practices validate the conclusion of Theme 4 that models that result in increased student engagement facilitate student achievement.

The fifth cross-case theme was that evaluation models promote an increase in student achievement when they incorporate student data. The use of data, which all participants (even those not required to use data) saw as helping drive increased student achievement, falls under Utility Standard 4. Recall that this fourth standard pertains to an evaluation system having clear and explicit criteria that directly relate to job expectations. As the standard's explanation makes clear, "[n]ot defining the criteria explicitly opens the opportunity for individual interpretation and jeopardizes the usefulness of results" (JCSEE, 2009, p. 85). This potential issue was seen when data was not required in that participants reported a generalized, individual understanding that student achievement was improving under their evaluation model, but could not point to specific evidence of that in particular classrooms. This leaves room for a rationale for Utility Standard 4 to be relevant: "Without explicit criteria, bias is more readily introduced into the evaluation" and this "diminishes the worth of the evaluation" (p. 85).

When an evaluator has a teacher's SLOs as part of the teacher's evaluation, there can be no individual interpretation as to whether that component was met; the desired student achievement either did or did not happen. This approach is supported by Marzano's developmental scale (2015), which includes teachers identifying a starting point, setting goals, and then tracking progress toward meeting those goals. By itself, data are neither good nor bad. It is how it is used that matters, and importantly, there are basic safeguards (e.g., not mandating

PERCEPTIONS OF TEACHER EVALUATION MODELS

universally high weights for data measures) that districts can take to minimize mistakes while still preserving the information the data provides (DiCarlo, 2012; Goe, 2013; Hanushek & Rivkin, 2010; Marshall, 2012). These principles were present in the two Cases that used data, and all of the participants agreed with the research suggesting that evaluation models need both specific and comprehensive growth focus to produce skilled teachers who are improving their teaching (Danielson, 2011; Marzano, 2015; Shaha et al., 2015; Tomlinson, 2014). Thus it seems similarly apparent that the conclusion of Theme 5 that evaluation models promote an increase in student achievement when they use student data is validated when the expectations from Utility Standard 4 that evaluation systems have and use explicit criteria are fulfilled.

Research Question 3: In what ways do teacher and administrator perceptions of the evaluation models influence the evaluation process and educational outcomes?

The sixth and final cross-case theme was a product of Research Question 3. The sixth theme concluded that evaluation models that are perceived positively are more likely to be used and, thus, will have a greater influence on process and outcomes. This theme implicated aspects of all six utility standards as each standard, in part, concerned enhancing the perception of those involved in the evaluation process. Utility Standard 1 was explained as being necessary so that evaluations were not seen as being used to comply with perfunctory bureaucratic requirements, to control or intimidate, or as a punitive system (JCSEE, 2009). Theme 6 spoke directly to the participants' beliefs that their models were seen as not doing those things. And this collective belief is a critical change from the view reported in research that pre-P.L. 90 that teachers tended to view evaluation systems as mostly punitive rather than fostering professional growth (Hanushek & Rivkin, 2010; Kane & Staiger, 2008). The second utility standard included in its rationale that evaluation systems had to guard against participant mistrust, which could trigger

PERCEPTIONS OF TEACHER EVALUATION MODELS

participants to fear information may be misused (JCSEE, 2009). The third standard's rationale focused even more directly on this notion of participant trust. It spoke directly about the acceptance of evaluations depending on the evaluatee's perceptions of the evaluator's qualifications, understanding of the evaluatee's position, and lack of bias. The importance of perceived fidelity of implementation was also stressed (2009). These standards match up clearly with the perceptions discussed in Theme 6 as they were the exact sorts of things uniformly discussed by the participants as impacting the buy-in for the models.

Similarly, the expectations pertaining to explicit criteria found in Utility Standard 4 and the creation and production of clear feedback found in Utility Standard 5 tie directly to Theme 6. Again, these are areas which, when not handled correctly, can cause intensely negative perceptions of an evaluation system. As these standards' rationales state, criteria that are not known, understood, and used consistently can result in non-uniform application with unfair results, and further "even the most accurate, timely [evaluation] will lose its ability to influence performance if not clearly understood by all users" (JCSEE, 2009, pp. 85, 90). The cross-case analysis made it clear that implementation of the models in a manner that overcame the concerns addressed by these standards was an important part in the positive perceptions of the evaluation models used at WHS, JHS, and LHS.

Finally, Utility Standard 6 is most directly related to Theme 6. The conclusion of this standard's rationale is that "[e]mpowering evaluatees to play a key role in their own growth helps build trust and respect between evaluators and evaluatees and promote staff morale" (p. 94). As with the previous standards, this was evidenced expressly in the sixth theme. Further, it is a notion supported by research that makes it clear that the most effective evaluation systems empower teachers to look positively at their practice to self-diagnose their identified areas of

PERCEPTIONS OF TEACHER EVALUATION MODELS

need and growth and allow them to improve their repertoire of skills (Frontier & Mielke, 2016).

Thus, this sixth theme adds to those before in terms of the Utility Standards supporting the conclusion iterated by the theme that positive perception of and buy-in for evaluation models make those models more effective both in their processes and their outcomes.

In summary, as was evident in the six themes that emerged from this study, the cross-case comparison of the three models produced evidence for best practice evaluations that lined up directly with the research on teacher evaluations and the Utility Standards, thus, providing the evidence that the models followed the research for development and process. There has been a definite need to change the past practice of teacher evaluation and move to a more effective model to improve overall teacher performance, which in turn would lead to higher levels of achievement by the students in schools (Alexander, 2016; Danielson & McGreal, 2000; Darling-Hammond, 2013; Marzano, 2012; Marzano & Toth, 2013; Tomlinson, 2014). The cross-case study of three models—RISE, TAP, and PAR—produced themes that demonstrated high expectations regarding professional growth and, thus, created a culture of evaluation in schools. The evaluation process described in the research is one that sees the balance between needing to know what goes on in the classroom and giving the teacher room to grow (Marshall, 1996; Marzano 2015; Peterson, 2000), and this was a balance referenced by the participants as being present in their evaluation models. Evaluation must increase the desire for teachers to engage in dialogue with supervisors and ensure that ongoing professional development be tied to goals. There must be work established to create clear and consistent definitions before any productive discourse can take place. To assess the quality of teaching practice, it is essential to define it. It is not sufficient to say, “I can’t define good teaching, but I know it when I see it” (Danielson, 2011, p. 36).

PERCEPTIONS OF TEACHER EVALUATION MODELS

The research pointed to the need in the state of Indiana for the development and use of new and improved teacher evaluation models. More specifically, new evaluation tools were needed that accomplished what effective evaluations should do—help to improve teaching practices and increase professionalism in the teaching profession, which will result in increased student achievement. The sorts of specific principles researchers agree on as being necessary for change (as outlined above) are the very practices that the participants in this study confirmed were happening within the three models they were using.

Thus, this study and the six themes that resulted from it has furthered the research in these areas, helped confirm the need for common expectations and “must haves” in all evaluations, and validated the previous conclusions about what really matters most. The literature review in this study connected well with the themes from the study and the two together showed that at least the principles stated in the themes would be needed for effective evaluations.

Unexpected findings

While the process of gathering data for my study went largely as expected, there were some findings that I found to be unexpected. Although I was not surprised by the fact that every school had its own teacher evaluation culture, I did not expect that each school felt so completely sure it had picked the right model and would not want to make any major evaluation model changes in the immediate future. I was surprised to hear no discussion at all related to thinking that the “grass was greener” with some other model than what was being used. Thus, the evaluators and practitioners across the state seem to value their ability to have a choice of which model to implement and to utilize, and perhaps this relates to the growth and achievement that the participants reported happening.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Second, one might be surprised that the participants were so generally positive about teacher evaluations systems, as evaluations tend to have a more negative connotation. I believe that was a product of professionals responding to something that is having a positive impact on them and their students coupled with the uniformly held belief that the evaluation systems were not administered in a punitive manner. My final cross-case theme speaks to this.

It is also possible the administrators who selected the teachers I was to interview naturally selected those teachers who were either highly effective and would be positively reinforced by the evaluation system or those who were known to value the evaluation system. Or, perhaps the administrators purposefully picked teachers who they felt would provide a positive slant to the evaluation model. Given this potential anomaly, I was intentional about asking probing questions to understand the aspects of the evaluation systems that were negative.

Finally, I did not expect the absolute level of professionalism and willingness to share with me that was exhibited by the teachers and administrators in this study. It was clear in my visits to the schools that all participants took all the questions seriously and prepared for every question ahead of time. Each of them in their own way was proud of the district they worked in, the model that had been chosen to use for teacher evaluations there, and the results their model was garnering. I thought that I might see participants burdened by participation and unwilling to contribute or attack issues deeply. I also thought I might see stress or a level of anxiety in discussing the evaluation process. However, I saw none of these things, and that was pleasantly surprising.

Conclusions

My perceived implications for practice, recommendations for further research, and concluding remarks follow.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Implications for Practice

The majority (seven of nine studies) of the research conducted post-P.L. 90 pertaining to teacher evaluations in Indiana that I found used a quantitative research methodology. Thus, this qualitative study added to this underrepresented category, and many different groups can benefit from the findings of this qualitative study.

This study supports the conclusions drawn from the six themes that emerged from the analysis of all relevant data. Those conclusions provided six specific areas, supported by research and the six Utility Standards, that administrators and teachers can use to look for and choose a teacher evaluation system that will support teacher effectiveness, raise student achievement accountability, and bring about desired professional growth.

The themes from this cross-case analysis of the three evaluation models align with many important research-based conclusions. For example, research would indicate (a) that educators have the professional responsibility and accountability to focus on teacher growth and effectiveness (Halverson & Clifford, 2006; Hill et al., 2012); and (b) teaching and learning are at the core of educational practice, and teacher quality is a very important school-level factor affecting student achievement (Goldhaber et al., 2013). The implication for practice is that school districts or Indiana Department of Education could look at my collected information of individual teacher and administrator perceptions to see that each evaluation model studied aligned with best practice evaluations strategies. Teachers, administrators, and district leaders can now determine whether each model—RISE, TAP, and PAR—was conducted with fidelity according to the rules of the evaluation model, methods, and tools used in evaluations, and whether teachers and administrators perceive them to have been effective. When I studied each of the evaluation models, I saw clear evidence that both the teachers and the administrators felt

PERCEPTIONS OF TEACHER EVALUATION MODELS

satisfied with their evaluation model. I still believe there is even more area to grow in using even more voices involved in the evaluation of teachers and the number of evaluations needing to increase.

Another implication was that data needs to be analyzed in order for accountability to be in place for both teachers and administrators. These three teaching evaluation models incorporated student achievement data. However, data were usually not used for decisions regarding employment and compensation in these districts. This is arguably a good thing given that research indicates that doing so could be harmful to positive teacher evaluation processes (Baker et al., 2010; Burris & Welner, 2011). Where evaluations go from this point in terms of the concepts of value-added and merit pay will be something that needs more research and cooperation between educators.

Another potential implication for practice is that teachers need to be even more involved in their own evaluation process. These new models encourage a systematic process of putting together a method of professional development for all participants that must be developed from the data and recommendations of the evaluations. This leads to having a model for growth and learning that the teachers would find professionally rewarding. There is a nearly uniform understanding that peer teaching and coaching models/programs are positive factors for growth and change in the evaluation of teachers. I would like to see all the models include a very clearly detailed outline regarding the importance that the role of consulting teachers can play in providing support to new and struggling teachers. The consulting teachers can offer this support and also conduct evaluations at the same time. For authentic feedback, teachers who take on this role must be trained and skilled to ask the tough question and to get to the heart of the matter for real buy-in and change to develop.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Additionally, this research will contribute to the existing body of literature focused on effective practices within each of the three teacher evaluation models—RISE, TAP, and PAR—themselves and will be beneficial given the potential for upcoming changes to Indiana teacher evaluation systems. This study contributed important evaluation ideas to educators because they will be able to make connections with the findings. This study provides educators an opportunity to match their own perceptions with that of the participants in this study.

Future elected officials, committee members, and leaders within the Indiana Department of Education can use the data from this study to determine if the current models are successful in their purpose. They can also use the data from this study on teacher and administrator perceptions to determine what changes need to be made to the models in order to improve teacher effectiveness. Personal connections can be made with this study because all educators are involved in the evaluation process.

Recommendations for further research

The recommendations from this qualitative case study include suggestions for teacher evaluation model design as well as for future research with regard to future-created teacher evaluation models throughout the State of Indiana. First, more research needs to be done to determine a level of consistency in the expectations for teachers and administrators. There will need to be a clear understanding of what is needed for quality teaching practice to be instituted in every classroom in the state of Indiana. Second, although much research has been done regarding the need to visit classrooms as frequently as possible, additional research about how frequently is optimal and how the optimal number might best happen in these models would be appropriate. Third, research should be conducted in the area of creating teacher evaluation models that meet the needs of teachers differently as related level of experience, i.e., new,

PERCEPTIONS OF TEACHER EVALUATION MODELS

middle, and veteran teacher differences. Finally, much could be learned from a state-to-state comparison of evaluation models and the perceptions provided from both teachers and administrators in different areas of the country.

Concluding remarks

In a little over five years' time, Indiana's schools have gone from a haphazard, lacking in uniform or research-based expectations system of evaluations (that was almost a system of non-evaluations) to a mandated system of evaluations. While that new system allows choice between certain evaluation models, all potential models are substantially more rigorous and prescribed than the earlier "non-system." This study sought to determine whether there is evidence that three of the different evaluation systems that a school district could choose under the new law—RISE, TAP, and PAR—were, in fact, doing what they were supposed to be doing: promoting professional growth and improving instructional practices, leading to increased student achievement. Six independent cross-case themes emerged from the research that all indicated that the research questions could be answered in the affirmative.

The six themes revealed that (a) evaluation models that include training, positive changes to instruction, and feedback contributed to professional growth; (b) models that include student-focused instruction and incorporate student data facilitate student achievement; (c) models that are perceived positively are more likely to be used and, thus, are more influential. These themes are directly aligned with the six Utility Standards that are the part of the 27 Personnel Evaluation Standards (JCSEE, 2009) used in this study and that should be considered further evidence that the research questions should be answered in the affirmative.

PERCEPTIONS OF TEACHER EVALUATION MODELS

REFERENCES

- Acheson, K. A., & Gall, M. D. (2003). *Clinical supervision and teacher development: Preservice and inservice applications*. Hoboken, NJ: John Wiley & Sons, Inc.
- Alexander, E. E. (2016). *Teacher Evaluation: The Relationship Between Performance Evaluation Ratings and Student Achievement*. (Doctoral dissertation), Liberty University, Lynchburg, VA.
- Almy, S. (2011). *Fair to everyone: Building the balanced teacher evaluations that educators and students deserve*. Education Trust: Washington, DC. Retrieved from <https://edtrust.org/resource/fair-to-everyone-building-the-balanced-teacher-evaluations-that-educators-and-students-deserve/>.
- Anderson, R. H. (1993). Clinical supervision: Its history and current context. *Clinical supervision: Coaching for higher performance*, 5-18.
- Andrews, K. A. (2015). *RISE: Evaluation of principals in Indiana* (Order No. 3745638). Available from ProQuest Dissertations & Theses A&I. (1756743307). Retrieved from <https://search-proquest-com.proxy.bsu.edu/docview/1756743307?accountid=8483>.
- Baker, B. D., Oluwole, J., & Green, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the race-to-the-top era. *Education Policy Analysis Archives*, 21(5).
- Baker, E. L., Barton, P.E., Darling-Hammond, L. Haertel, E., Ladd, H.F., Linn, R. L., Ravitch, D.... Shepard, L.A. (2010). Problems with the use of student test scores to evaluate teachers. Retrieved from <http://www.epi.org/publication/bp278/>.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Bambrick-Santoyo, P. (2012). Beyond the scoreboard. *Educational Leadership*, 70(3), 26-30.

Retrieved from <http://www.ascd.org/publications/educational-leadership/nov12/vol70/num03/Beyond-the-Scoreboard.aspx>.

Barnes, W. B., & Slate, J. R. (2013). College-readiness is not one-size-fits-all. *Current Issues in Education*, 16(1).

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, 13(4), 544-559.

Blume, H. (2012). Lawsuit takes aim at California's legal protections for teachers. Los Angeles Times. Retrieved from <http://articles.latimes.com/2012/may/16/local/la-me-0516-laUSD-teachers-20120516>

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theory and methods* (3rd ed.). Boston, MA: Allyn & Bacon.

Boyland, L., Harvey, M.W., Quick, M., & Choi, Y. (2014). Teacher evaluation in Indiana: Principals' use and perception of field-based practices. *International Journal of Educational Reform*, 23, 270-293.

Burke, L. M., & Marshall, J. A. (2010). Why National Standards Won't Fix American Education: Misalignment of Power and Incentives. Background. No. 2413. *Heritage Foundation*.

Burris, C. C. & Welner, K.G. (2011). Conversations with Arne Duncan: Offering advice on educator evaluations. *The Phi Delta Kappan*, 93(1): 38-41.

Civic Impulse. (2016). H.R. 1532 – 112h Congress: Race to the Top Act of 2011. Retrieved from <https://www.govtrack.us/congress/bills/112/hr1532>.

Cogan, M. (1972). *Clinical supervision*. Boston, MA: Houghton Mifflin.

PERCEPTIONS OF TEACHER EVALUATION MODELS

- Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher*, 45(6), 378-387.
- Corbin, J., & Strauss, A. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.
- Crawford, J. (2004). *Educating English learners: Language diversity in the classroom*. Bilingual Education Serv.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.
- Cubberley, E. P. (1929). *Public School Administration...Revised and Enlarged Edition*. Houghton Mifflin Company.
- Danielson, C. (1996). *A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35-39.
Retrieved from <http://www.ascd.org/publications/educational-leadership/dec10/vol68/num04/Evaluations-That-Help-Teachers-Learn.aspx>.
- Danielson, C. (2012). Observing classroom practice. *Educational Leadership*, 70(3), 32-37.
Retrieved from <http://www.ascd.org/publications/educational-leadership/nov12/vol70/num03/Observing-Classroom-Practice.aspx>.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (1996). The right to learn and the advancement of teaching: Research, policy, and practice for democratic education. *Educational Researcher*, 25(6), 5-17.

PERCEPTIONS OF TEACHER EVALUATION MODELS

- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York, NY: Teachers College Press.
- Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher*, 44(2), 132-137. <http://dx.doi.org/10.3102/0013189X15575346>.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, 93(6), 8-15.
<http://dx.doi.org/10.1177/003172171209300603>.
- Darling-Hammond, L., Barron, B., Pearson, P. D., Schoenfeld, A. H., Stage, E. K., Zimmerman, T. D., & Tilson, J. L. (2015). *Powerful learning: What we know about teaching for understanding*. John Wiley & Sons.
- Dewey, J. (1938). *Experience and education*. New York, NY: Kappa Delta Pi.
- DiCarlo, M. (2012). How to use value-added measures right. *Educational Leadership*, 70(3), 38-42. Retrieved from
http://www.ascd.org/publications/educational_leadership/nov12/vol70/num03/How_to_Use_Value-Added_Measures_Right.aspx.
- Dillon, N. (2013). How are your teachers and principals doing? *American School Board Journal*, 13(8), 30-31. Retrieved from <http://www.asbj.com>.
- Duffy, F. M. (1998). The ideology of supervision. In G. Firth & E. Pajak (Eds.) *Handbook of Research on School Supervision* (pp. 181-199). New York: Macmillan.
- Eberline, A. D. (2016). *Perceptions of and experiences with the indiana teacher evaluation system in physical education* (Doctoral dissertation). Retrieved from <https://search-proquest-com.proxy.bsu.edu/docview/1849004951?accountid=8483>. (10170584).
- Eisner, E. W. (2017). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. Teachers College Press.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Elementary and Secondary Education Act of 1965, 20 U.S.C. 2701 *et seq.*

Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, 17(1), 101-128.

Fisher, D., & Frey, N. (2014). Using teacher learning walks to improve instruction. *Principal Leadership*, 58-61. Retrieved from http://library.standrews-de.org/lists/cap-articles/principle-leadership_teacher-walks_2014_01.pdf.

Ford, T. G., Van Sickle, M. E., & Fazio-Brunson, M. (2016). The Role of “Informational Significance” in Shaping Louisiana Elementary Teachers’ Use of High-Stakes Teacher Evaluation Data for Instructional Decision-Making. In *Student Growth Measures in Policy and Practice* (pp. 117-135). Palgrave Macmillan, New York.

Frontier, T., & Mielke, P. (2016). *Making teachers better, not bitter: Balancing evaluation, supervision, and reflection for professional growth*. ASCD.

Gallagher, H. A. (2004). Vaughn elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education*, 79(4), 79-107. http://dx.doi.org/10.1207/s15327930pje7904_5.

Gitlin, A., & Smyth, J. (1989). *Teacher evaluation: Educative alternatives.*, Philadelphia, PA: Falmer Press, Taylor & Francis, Inc.

Goe, L. (2013). Can teacher evaluation improve teaching? *Principal Leadership*, 13(7), 24-29. Retrieved from <http://www.nassp.org/Content/158/PLmar13+goe.pdf>.

Goe, L., Biggers, K., & Croft, A. (2012). Linking Teacher Evaluation to Professional Development: Focusing on Improving Teaching and Learning. Research & Policy Brief. *National Comprehensive Center for Teacher Quality*.

PERCEPTIONS OF TEACHER EVALUATION MODELS

- Goldhaber, D. D., Goldschmidt, P., & Tseng, F. (2013). Teacher value-added at the high-school level: Different models, different answers?. *Educational Evaluation and Policy Analysis*, 35(2), 220-236.
- Goldhammer, R. (1969). *Clinical supervision*. New York, NY: Holt, Rinehart, & Winston.
- Goldring, E., Grissom, J. A., Rubin, M., Neumerski C. M., Cannata, M., Drake, T., & Schuermann, P. (2015). Make room value added: Principals' human capital decisions and the emergence of teacher observation data. *Educational Researcher*, 44(2), 96-104.
<http://dx.doi.org/10.3102/0013189X15575031>.
- Goldstein, J. (2008). Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review. *American Journal of Education*, 113, 479-508.
<http://dx.doi.org/10.1086/512741>.
- Guskey, T. R. (2000). *Evaluating professional development*. Corwin Press.
- Guthrie, J. W., & Springer, M. G. (2004). A nation at risk revisited: Did "wrong" reasoning result in "right" results? At what cost? *Peabody Journal of Education*, 79(1), 7-35.
- Hall, T. (2002). *Differentiated instruction*. Wakefield, MA: National Center on Accessing the General Curriculum, CAST, U.S. Office of Special Education Programs.
- Halverson, R. R., & Clifford, M. A. (2006). Evaluation in the wild: A distributed cognition perspective on teacher assessment. *Educational Administration Quarterly*, 42(4), 578-619.
- Hanushek, E., & Rivkin, S. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100, 267-271.
<http://dx.doi.org/10.1257/aer.10.2.267>.

PERCEPTIONS OF TEACHER EVALUATION MODELS

- Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal*, 51(1), 73-112. <http://dx.doi.org/10.3102/0002831213517130>.
- Hattie, J., & Timperly, H. (2007). The power of feedback: Review of educational research. *Review of Educational Research*, 77(1), 81-112. <http://dx.doi.org/10.3102/00346543029847>.
- Hazi, H. M., & Rucinski, D. A. (2009). Teacher evaluation as a policy target for improved student learning: A fifty-state review of statute and regulatory action since NCLB. *Educational Policy Analysis Archives*, 17(5). <http://dx.doi.org/10.14507/epaa.v17.n5.2009>.
- Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher*, 41(2), 56-64.
- Hoy, W. K. (2010). *Quantitative research in education: A primer*. Thousand Oaks, CA: Sage.
- Humphrey, D. C., Koppich, J. E., Bland, J. A., & Bosetti, K. R. (2011). Peer review: Getting serious about teacher support and evaluation. *Menlo Park: SRI International*. Retrieved August, 7, 2012.
- Hunter, M. C. (1982). *Mastery teaching*, El Segundo, CA: Tip Publications.
- Hunter, R. (2004). *Madeline Hunter's mastery teaching: Increasing instructional effectiveness in elementary and secondary schools*. Corwin Press.
- Hursh, D. (2007). Exacerbating inequality: The failed promise of the No Child Left Behind Act. *Race Ethnicity and Education*, 10(3), 295-308.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Indiana Code § 20-28-11.5

Indiana Department of Education (IDOE). (2011). *Changing the evaluation process in Indiana*.

Retrieved from <http://www.doe.in.gov/sites/default/files/evaluations/rise-handbook-principals.pdf>

Indiana Department of Education (IDOE). (2012). *RISE Indiana teacher effectiveness rubric*.

Retrieved from <http://www.doe.in.gov/sites/default/files/evaluations/rise-handbook-principals.pdf>

Johnson, M. A., & Stephens, M. L. (2012). Race to the top and the exclusion of welfare recipients from educational policy discourse. *Adult Learning*, 23(4), 188-195.

<http://dx.doi.org/10.1177/1045159512458261>.

Johnson, S. M., & Fiarman, S. E. (2012). The potential of peer review. *Educational Leadership*,

70(3), 20-25. Retrieve from <http://www.ascd.org/publications/educational-leadership/nov12/vol70/num03/The-Potential-of-Peer-Review.aspx>.

Joint Committee on Standards for Educational Evaluation (JCSEE). (2009). *The personnel evaluation standards* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Kane, T., & Cantrell, S. (2012). *Learning about teaching: Initial findings from the measures of effective teaching project*. Seattle, WA: Bill and Melinda Gates Foundation.

Kane, T., & Staiger, D. (2008). *Estimating impacts on student achievement: An experimental evaluation* (NBER Working Paper No. 14607). Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w14607>.

Kellaghan, T., & Stufflebeam, D. L. (Eds.). (2012). *International handbook of educational evaluation: Part one: Perspectives/Part two: Practice* (Vol. 9). Springer Science & Business Media.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Keller, B. (2006). Teachers to conduct peer reviews in Chicago. *Education Week*, 25(20), 8.

Retrieved from <http://www.edweek.org/ew/articles/2006/01/25/20chicago.h25.html>

Knoeppel, R. C., & Rinehart, J. S. (2008). Student achievement and principal quality: Explaining the relationship. *Journal of School Leadership*, 18(5), 501-527.

Koppich, J. E., & Humphrey, D. C. (2011). Getting serious about teacher evaluation.

Education Week, 1, 18. Retrieved from

<http://www.edweek.org/ew/articles/2011/10/12/07koppich.h31.html>.

Lagemann, E. C. (2002). *An elusive science: The troubling history of education research*.

University of Chicago Press.

Laurillard, D. (2013). *Teaching as a design science: Building pedagogical patterns for learning and technology*. Routledge.

Lazaridou, A. & Iordanides, G. (2011). The principal's role in achieving school effectiveness.

International Studies in Educational Administration, 39(3), 3-19.

Lee, J., & Orfield, G. (2006). *Tracking achievement gaps and assessing the impact of NCLB on the gaps: An in-depth look into national and state reading and math outcomes*.

Cambridge, MA: Harvard Civil Rights Project.

Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning*. New York, NY: Wallace Foundation. Retrieved

from <http://www.wallacefoundation.org/knowledge-center/Documents/How-Leadership-Influences-Student-Learning.pdf>.

Lichtman, M. (2013). *Qualitative research for the social sciences*. Thousand Oaks, CA: Sage.

Looney, J. (2011). Developing high-quality teachers: Teacher evaluation for improvement.

European Journal of Education, 46, 440-455. <http://dx.doi.org/10.1111/j.1465-3435.2011.01492.x>.

PERCEPTIONS OF TEACHER EVALUATION MODELS

- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In *The scholarship of teaching and learning in higher education: An evidence-based perspective* (319-383). Springer Netherlands.
- Marshall, K. (1996). How I confronted HSPS (hyperactive superficial principal syndrome) and began to deal with the heart of the matter. *The Phi Delta Kappan*, 77, 336-345. Retrieved from <http://www.ocmboces.org/tfiles/folder1602/HowIConfrontedHSPS.pdf>.
- Marshall, K. (2012). Fine-tuning teacher evaluation. *Educational Leadership*, 70(3), 50-53.
- Marshall, C., & Rossman, G. B. (2011). *Designing qualitative research* (5th ed.). Thousand Oaks, CA: Sage.
- Marzano, R. J. (2012). The Two Purposes of Teacher Evaluation. *Educational Leadership*, 70(3), 14-19.
- Marzano, R. J. (2015). *Marzano teacher evaluation model: A cause-and-effect relationship to student achievement*. Blairsville, PA: Learning Sciences International.
- Marzano, R. J., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Mathis, W. J. (2011). Race to the Top: An example of belief-dependent reality. A response to "Race to the Top leaves children and future citizens behind". *Democracy and Education*, 19(2), Article 6.
- McGuinn, P. (2012). Stimulating reform: Race to the Top, competitive grants and the Obama education agenda. *Educational Policy*, 26(1), 136-159.
- Mears, C.L. (2009). *Interviewing for education and social science research: The gateway method*. New Your, NY: Palgrace MacMillan.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Merriam, S.B. (2009). *Qualitative research*. San Francisco: Jossey-Bass.

Morgan, G. B., Hodge, K. J., Trepinski, T. M., & Anderson, L. W. (2014). The stability of teacher performance and effectiveness: Implications for policies concerning teacher evaluation. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, 22.

Munson, B. R. (1998). Peers observing peers: The better way to observe teachers. *Contemporary Education*, 69(2), 108-110. Retrieved from <http://eric.ed.gov/?id=EJ564693>.

National Commission on Excellence in Education (NCEE). (1983). *A nation at risk: The imperative for educational reform*. Washington, DC. Retrieved from <http://www2.ed.gov/pubs/NatAtRisk/index.html>.

National Conference of State Legislatures (NCSL). (2005). *Task force on No Child Left Behind Final Report*. Denver.

National Governors Association (NGA). (2011). Governors play role in preparing principals to conduct teacher evaluations. Retrieved from http://www.nga.org/cms/home/news-room/news-releases/page_2011/col2-content/main-content-list/governors-play-role-in-preparing.html.

Natriello, G. (1990). Intended and unintended consequences: Purposes and effects of teacher evaluation. *The new handbook of teacher evaluation*: Newbury Park, CA: Sage.

Newman, J. W. (1992). Architect of the New Educational Hierarchy: Ellwood P. Cubberley. *Teaching Education*, 4(2), 161-168.

Neville, R. F., & Garmen, N. B. (1998). The philosophical perspective on supervision. *Handbook of Research on School Supervision*, 200-241.

No Child Left Behind Act of 2001, Pub. L. No. 107-110 (2002).

PERCEPTIONS OF TEACHER EVALUATION MODELS

- Nolan Jr. J., & Hoover, L.A. (2011). *Teacher supervision and evaluation*. John Wiley and Sons.
- Oliva, M., Mathers. C., & Laine, S. (2009). Effective evaluation. *Principal Leadership*, 9(7),16-21.
- Onosko, J. (2011). Race to the Top leaves children and future citizens behind: The devastating effects of centralization, standardization, and high stakes accountability. *Democracy and Education*, 19(2), 1.
- Paige, R. (2002). *Meeting the highly qualified teachers challenge: The secretary's annual report on teacher quality*. Washington DC: U.S. Department of Education.
- Papay, J. P., & Johnson, S. M. (2012). Is PAR a good investment? Understanding the costs and benefits of teacher peer assistance and review programs. *Educational Policy*, 26, 696-729. <http://dx.doi.org/10.1177/0895904811417584>.
- Patton, M.Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices*. Corwin Press.
- Pieczura, M. (2012). Weighing the pros and cons of TAP. *Educational Leadership*, 70(3), 70-72. Retrieved from <http://www.ascd.org/publications/educational-leadership/nov12/vol70/num03/Weighing-the-Pros-and-Cons-of-TAP.aspx>.
- Popham, J. (2013). On serving two masters: Formative and summative teacher evaluation. *Principal Leadership*, 13(7), 18-22. Retrieved from <http://eric.ed.gov/?id=EJ101642>. 1
- Protheroe, N. (2002). Improving instruction through teacher observation. *Principal*, 82(1), 48-51. Retrieved from <http://eric.ed.gov/?id=EJ650423>.

PERCEPTIONS OF TEACHER EVALUATION MODELS

- Protsik, J. (1995). History of Teacher Pay and Incentive Reforms. *Journal of School Leadership*, 6(3), 265-89.
- Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.
- Roberts, C. M. (2010). *The dissertation journey: A practical and comprehensive guide to planning, writing, and defending your dissertation*. Thousand Oaks, CA: Corwin.
- Saldaña, J. (2009). *The coding manual for qualitative researchers*. Los Angeles, CA: Sage.
- Sargent, M. S. (2014). *An investigation of research-based teaching practices through the teacher evaluations in indiana public schools* (Doctoral dissertation). Indiana State University, Terre Haute, IN. (3616908).
- Sartain, L. Stoelinga, S. R., & Brown, E. R. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation*. Chicago, IL: Consortium on Chicago School Research.
- Saultz, A., & Saultz, J. W. (2017). Measuring outcomes: Lessons from the world of public education. *The Annals of Family Medicine*, 15(1), 71-76.
- Sawchuk, S. (2010). New teacher distribution methods hold promise. *Education Week*, 29(35), 1-16. Retrieved from <http://www.edweek.org/ew/articles/2010/06/10/35distribute.h29.html>.
- Scannella, A., & McCarthy, S. (2014). Adversity opportunity? *Principal Leadership*, 14(5), 52-55. Retrieved from <http://www.nassp.org>.
- Schachter, R. (2012). Brave new world of teacher evaluation. *District Administration*, 48(10), 43-47. Retrieved from <https://www.districtadministration.com/article/brave-new-world-teacher-evaluation>.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (4th ed.). New York, NY: Teachers College Press.

Selwyn, D. (2007). Highly quantified teachers NCLB and teacher education. *Journal of Teacher Education*, 58(2), 124-137.

Shaha, S. H., Glassett, K. F., & Copas, A. (2015). The impact of teacher observations with coordinated professional development on student performance: A 27-state program evaluation. *Journal of College Teaching & Learning (TLC)*, 12(1), 55-64.
<http://dx.doi.org/10.19030/tlc.v12i1.9070>.

Smith, W. F., & Andrews, R.L. (1989). *Instructional Leadership: How Principals Make a Difference*, Alexandria, VA: Association for Supervision and Curriculum Development.

Solmon, L., White, J. T., Cohen, D., & Woo, D. (2007). *The effectiveness of the teacher advancement program*. Santa Monica, CA: National Institute for Excellence in Teaching.

Spring, J. (2010). *The politics of American education*. New York, NY: Routledge.

Stake, R. (2005). Qualitative case studies. In N. Denzin & Y. Lincoln (Eds.). *The Sage handbook of qualitative research* (pp. 443-466). Thousand Oaks, CA: Sage.

Steinberg, M. P., & Donaldson, M. L. (2016). The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era. *Education Finance and Policy*, 11(3), 340-359.

Stronge, J. H., & Ostrander, L. P. (2006). Client surveys in teacher evaluation. In *Evaluating teaching: A guide to current thinking and best practice* (2nd ed.) (pp. 125-151). Thousand Oaks, CA: Corwin Press.

Stonge, J. H., Tucker, P. D., & Hindman, J. L. (2004). *Handbook for qualities of effective teachers*. ACSD.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Sullivan, S., & Glanz, J. (2005). *Supervision that improves teaching: Strategies and techniques*. Corwin Press.

Superfine, B. M. (2016). Legal issues and implications of teacher evaluation systems. In Hewett, K. K., Amrein-Beardsley, A. (Eds.), *Student growth measures: Where policy meets practice*. New York, NY: Palgrave Macmillan.

Superfine, B. M., Gottlieb, J. J., & Smylie, M. A. (2012). The expanding federal role in teacher workforce policy. *Educational Policy*, 26(1), 58-78.

Taylor, F. W. (2004). *Scientific management*. London, England: Routledge.

Taylor, E. S., & Tyler, J. H. (2012). Can teacher evaluation improve teaching? *Education Next*, 12(4).

Thorne, S. (2016). *Interpretive description: Qualitative research for applied practice* (Vol. 2). Routledge.

Timperley, H. (2011). *Realizing the power of professional learning*. United Kingdom: McGraw-Hill Education.

Toch, T., & Rothman, R. (2008). Rush to judgment: Teacher evaluation in public education. *Education Sector*.

Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners*. ASCD.

U.S. Department of Education. 2009. Race to the Top Program Executive Summary.

U.S. Department of Education (2012). Department releases state reports profiling first-year progress under Race to the Top. Retrieved from <http://www.ed.gov/news/press-releases/departments-releases-state-report>

PERCEPTIONS OF TEACHER EVALUATION MODELS

- Watts, G. W. (2016). An overview of the issues in evaluating special educators: Current challenges and recent developments. 4(1), 46-53.
- Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. *Education Digest: Essential Readings Condensed for Quick Review*, 75(2), 31-35.
- Weiss, R. (1994). *Learning from strangers: The art and method of qualitative interview studies*. New York, NY: Free Press.
- Wetzel, W. A. (1932). Characteristics of pupil leaders. *The School Review*, 40(7), 532-534.
- Whiteman, R. S., Shi, D., & Plucker, J. A. (2011). Revamping the teacher evaluation process. *Education Policy Brief*, 9(4), 1-24.
- Wise, A. E. (1984). *Teacher evaluation: A study of effective practices*. Santa Monica: CA: Rand.
- Yin, R.K. (2013). *Case study research: Design and Methods*. Sage Publications.
- Zatynski, M. (2012). Revamping teacher evaluation. *Principal*, 91(5), 22-27. Retrieved from <http://www.naesp.org/sites/default/files/MK12%20Zatynski.pdf>.
- Zeichner, K. (2011). Assessing state and federal policies to evaluate the quality of teacher preparation programs. *Teacher education policy in the United States: Issues and tensions in an era of evolving expectations*, 76-105. New York, NY: Routledge.
- Zepeda, S. J. (2014). *The principal as instructional leader: A handbook for supervisors*. Routledge.
- Zhao, Y. (2009). *Catching up or leading the way: American Education in the Age of Globalization*. ASCD.

PERCEPTIONS OF TEACHER EVALUATION MODELS

APPENDIX A: INTERVIEW QUESTIONS FOR ALL PARTICIPANTS

Process:

- An appointment will be made with each of the two teachers and administrator for 60 minutes.
- The interview will be audio recorded for transcription.
- Questions for clarity may be added.
- Explanation of the question should be provided if requested.

Script: This interview is intended to provide me with information about the teacher evaluation model in your school district. If you would like more explanation of a question, feel free to ask at any time. If you would prefer to not answer a question, please let me know that as well. For the purpose of this interview, please think about the teacher evaluations you have had or conducted in your educational setting. Please remember your answers will not be shared with anyone within the three high schools nor with anyone at the school district.

Overarching question: From the viewpoints of both teachers and administrators, do the three teacher evaluation models—RISE, TAP, or PAR—actually in practice, cause teacher growth and improve student achievement?

Preliminary information: Please describe what you do here, your length of service, and any previous service as an educator.

Research Question 1: What characteristics of the teacher evaluation models do teachers and administrators perceive to be the most effective in promoting professional growth?

- a. Would you please walk me through the teacher evaluation model used in your district? Have any forms or tools been created that help provide a robust evaluation with feedback to help the evaluation process? Please share with me what you use.
- b. Since the implementation of the teacher evaluation model, describe your feelings about your professional growth as a teacher or administrator. What has changed and what has not?
- c. To what extent has the teacher evaluation model changed the way you deliver instruction in the classroom as a teacher and the evaluation practices done by the administrator? If no changes have occurred, why do you think not?

PERCEPTIONS OF TEACHER EVALUATION MODELS

- d. Describe the components of the teacher evaluation model that have provided you as a teacher or administrator the best opportunities for growth. Could you provide a couple of examples?
- e. Define the types of training provided to teachers and administrators during the implementation of the teacher evaluation and the on-going implementation phases.
- f. Have there been efforts to improve the implementation of teacher evaluations, and if so, could you describe those efforts in detail?

Research Question 2: What characteristics of the teacher evaluation models do teachers and administrators perceive to be the most effective in promoting student achievement?

- a. Tell me about your predominant instructional strategies. Have they changed or evolved over the years using the teacher evaluation model?
- b. To what extent has the teacher evaluation system required you to incorporate student achievement or student data as a teacher or administrator? Describe the process the district uses for the inclusion of data for student achievement.
- c. Characterize challenges to the process of implementing teacher evaluations from either the teacher or administrator point of view.
- d. What specific characteristics in the model have helped improve student achievement? Please provide examples.

Research Question 3: In what ways do teacher and administrator perceptions of the evaluation models influence the evaluation process and educational outcomes?

- a. In what ways does the teacher evaluation system align with the beliefs of a good lesson and a well-run classroom? Describe how the model is the model helping or not helping you in your education processes and outcomes.
- b. Describe the changes made in the classroom by the teacher or the administrative practices to improve instruction.
- c. Describe the process of reflection and dialogue used between teacher and administrator.
- d. What is your perception of how other teachers and administrators feel about the evaluation model?

PERCEPTIONS OF TEACHER EVALUATION MODELS

APPENDIX B: RECRUITMENT SCRIPT FOR POTENTIAL TEACHER PARTICIPANTS

Recruitment Script for Teacher
Introductory Script Phone Message

On the phone:

“Hello, my name is Bret Daghe. I am a doctoral student in Ball State University’s Department of Educational Leadership. I received your name from [insert administrator’s name]. I am conducting a research study regarding a comparison of Indiana administrator and teacher perceptions of teacher evaluation models, and I have received approval from [insert superintendent’s name] to conduct my research at [insert school name]. I am hoping that you would be willing to be a participant in my study and let me interview you. It should take about sixty minutes to complete the interview and can be done at your convenience.

[If the teacher indicates interest/willingness, we will set up a date and time then.]

[If the teacher seems uncertain:] “I certainly understand your wanting to think about it. Why don’t I let you do that; would you consider getting back to me by [some period 3-5 days away].”

[In either of those scenarios, I will inform the teacher of the following information:]

“For form’s sake, let me tell you that to be eligible to participate in this study, you must be a teacher with [which ever criteria of years teaching are applicable - five or fewer years of experience at that school or six to twenty years] of experience at [insert school name]. Your identity as a participant will be anonymous. Further, for questions about your rights as a participant, you may contact Director, Office of Research Integrity, Ball State University, Muncie, IN 47306, (765) 285-5070, irb@bsu.edu.”

“To reiterate, I have you scheduled for an interview on _____. If you have any questions, I can be reached at 317-372-8405 or bldaghe@bsu.edu.”

[If the teacher is not interested:] “I understand and appreciate you being candid with me now. Thank you for your time, and I hope you have a great fall semester.”

Bret L. Daghe
Doctoral Candidate
Department of Educational Leadership
Ball State University
Teachers College 925, Muncie, IN 47306
Telephone: (317) 372-8405
Email: bldaghe@bsu.edu

PERCEPTIONS OF TEACHER EVALUATION MODELS

APPENDIX C: RECRUITMENT SCRIPT FOR POTENTIAL ADMINISTRATOR

PARTICIPANTS

Recruitment Script for Administrator Introductory Script Phone Message

On the phone:

“Hello, my name is Bret Daghe. I am a doctoral student in Ball State University’s Department of Educational Leadership. I received your name from [insert superintendents’ name]. I am conducting a research study regarding a comparison of Indiana administrator and teacher perceptions of teacher evaluation models, and I have received approval from [insert superintendent’s name] to conduct my research at [insert school name]. I am hoping that you would be willing to be a participant in my study and let me interview you. It should take about sixty minutes to complete the interview and can be done at your convenience.

[If the administrator indicates interest/willingness, we will set up a date and time then.]

[If the administrator seems uncertain:] “I certainly understand your wanting to think about it. Why don’t I let you do that; would you consider getting back to me by [some period 3-5 days away].”

[In either of those scenarios, I will inform the administrator of the following information:]
“Let me also tell you that to be eligible to participate in this study, you must be the administrator with the greatest knowledge and expertise at [insert school name] with the [insert evaluation model name] model. Your identity as a participant will be anonymous. Further, for questions about your rights as a participant, you may contact Director, Office of Research Integrity, Ball State University, Muncie, IN 47306, (765) 285-5070, irb@bsu.edu.”

“To reiterate, I have you scheduled for an interview on _____. If you have any questions, I can be reached at 317-372-8405 or bldaghe@bsu.edu.”

[If the administrator is not interested:] “I understand and appreciate you being candid with me now. Thank you for your time, and I hope you have a great fall semester.”

Bret L. Daghe
Doctoral Candidate
Department of Educational Leadership
Ball State University
Teachers College 925, Muncie, IN 47306
Telephone: (317) 372-8405
Email: bldaghe@bsu.edu

PERCEPTIONS OF TEACHER EVALUATION MODELS

APPENDIX D: INFORMED CONSENT FOR TEACHER PARTICIPANTS

A COMPARISON OF INDIANA ADMINISTRATOR AND TEACHER PERCEPTIONS OF TEACHER EVALUATION MODELS: TEACHER CONSENT FORM

Study Title: A Comparison of Indiana Administrator and Teacher Perceptions of Teacher Evaluation Models

Study Purpose and Rationale: The purpose of the study is to analyze three High-Stakes Accountability Models--RISE Evaluation and Development Systems (RISE), Teacher Advancement Program (TAP), and Peer Assistance and Review (PAR). Data will be collected on teacher and administrator perceptions concerning: (a) the general methodology or structure of the three evaluation models; (b) the specific practices for gathering and compiling the raw data in each of the three models (i.e., the forms used); and (c) the potential impact of each model on teacher effectiveness, student achievement, and professional growth. The rationale of this qualitative study is critical for both teachers and administrators because evaluations purport to measure teacher effectiveness and teacher effectiveness is thought to impact student learning. Research into the methodology, the specific practices, and the outcomes of the three different models for evaluation can be used to improve the overall process. The research will contribute to the existing body of literature focused on effective practices within each of the three models and will be beneficial given the ongoing evaluation of the teacher evaluation system in Indiana. I am unaware of previous work that captures teacher perceptions in addition to administrator perceptions regarding the three evaluation models.

Inclusion/Exclusion Criteria: At each school, you will be a teacher with five or fewer years of teaching experience, or a teacher with six to twenty years of teaching experience. Your age is not the relevant criteria, the years of experience is; however, you will need to be over 21 years of age. There are no exclusion criteria for you other than falling outside of the years of experience needed.

Participation Procedures and Duration: Your interviews will last approximately one hour and will take place at a time and place convenient to you.

Audio Tapes: For purposes of accuracy, with your permission, the interviews will be audio recorded. Your name used on the audiotape will be changed to pseudonyms when the recordings are transcribed. Once transcribed and the research findings are complete, the recordings will be deleted. The transcripts will be kept for 3 years.

Disclosure of Alternative Procedures: This is not applicable to you for this research study.

Data Confidentiality or Anonymity: All data will be maintained as confidential (if collecting identifiable data, i.e., audio/video recordings) and no identifying information such as your name will appear in any publication or presentation of the data. The data will be confidential. I will be conducting in person interviews with you; however, I will use pseudonyms, and only I will know the match between the pseudonym and your identity and the schools.

Storage of Data and Data Retention Period: Paper data will be stored in a locked filing cabinet in the researcher's office for three years and then will be shredded. The data will also be entered into a software program and stored on the researcher's password-protected computer for three years and then deleted. Only members of the research team will have access to the data.

Risks or Discomforts: There are no anticipated risks or discomforts for you with this study.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Who to Contact Should You Experience Any Negative Effects from Participating in this Study:

Should you experience negative effects, please contact a state counseling services provider for an appointment or for a recommendation for services with another provider.

Benefits: There are no perceived benefits for your participation in this study.

Voluntary Participation: Your participation in this study is completely voluntary and you are free to withdraw your permission at any time for any reason without penalty or prejudice from the investigator. Please feel free to ask any questions of the investigator before signing this form and at any time during the study.

IRB Contact Information: For your rights as a participant, you may contact the following: For questions about your rights as a participant, please contact the Director, Office of Research Integrity, Ball State University, Muncie, IN 47306, (765) 285-5070 or at irb@bsu.edu.

Study Title: A Comparison of Indiana Administrator and Teacher Perceptions of Teacher Evaluation Models

Consent

I, _____, agree to participate in this research project entitled, A Comparison of Indiana Administrator and Teacher Perceptions of Teacher Evaluation Models. I have had the study explained to me and my questions have been answered to my satisfaction. I have read the description of this project and give my consent to participate. I understand that I will receive a copy of this informed consent form to keep for future reference.

To the best of my knowledge, I meet the inclusion/exclusion criteria for participation (described on the previous page) in this study.

Participant's Signature

Date

Researcher Contact Information

Principal Investigator:

Bret L. Daghe, Doctoral Student
Department of Educational Leadership
Ball State University
Muncie, IN 47306
Telephone: (765) 285-8490
Email: bldaghe@bsu.edu

Faculty Supervisor:

Dr. Kendra Lowery
Department of Educational Leadership
Ball State University
Muncie, IN 47306
Telephone: (765) 285-8490
Email: kplowery@bsu.edu

PERCEPTIONS OF TEACHER EVALUATION MODELS

APPENDIX E: INFORMED CONSENT FOR ADMINISTRATOR PARTICIPANTS

A COMPARISON OF INDIANA ADMINISTRATOR AND TEACHER PERCEPTIONS OF TEACHER EVALUATION MODELS: ADMINISTRATOR CONSENT FORM

Study Title: A Comparison of Indiana Administrator and Teacher Perceptions of Teacher Evaluation Models

Study Purpose and Rationale: The purpose of the study is to analyze three High-Stakes Accountability Models--RISE Evaluation and Development Systems (RISE), Teacher Advancement Program (TAP), and Peer Assistance and Review (PAR). Data will be collected on teacher and administrator perceptions concerning: (a) the general methodology or structure of the three evaluation models; (b) the specific practices for gathering and compiling the raw data in each of the three models (i.e., the forms used); and (c) the potential impact of each model on teacher effectiveness, student achievement, and professional growth. The rationale of this qualitative study is critical for both teachers and administrators because evaluations purport to measure teacher effectiveness and teacher effectiveness is thought to impact student learning. Research into the methodology, the specific practices, and the outcomes of the three different models for evaluation can be used to improve the overall process. The research will contribute to the existing body of literature focused on effective practices within each of the three models and will be beneficial given the ongoing evaluation of the teacher evaluation system in Indiana. I am unaware of previous work that captures teacher perceptions in addition to administrator perceptions regarding the three evaluation models.

Inclusion/Exclusion Criteria: At each school, you will be the administrator interviewed because you have the most knowledge about and experience with the evaluation model used by the school. Your age is not the relevant criteria, the years of experience is; however, you will need to be over 21 years of age. There are no exclusion criteria for you other than falling outside of the years of experience needed.

Participation Procedures and Duration: Your interviews will last approximately sixty minutes and will take place at a time and place convenient to you.

Audio or Video Tapes (if applicable): For purposes of accuracy, with your permission, the interviews will be audio recorded. Your name used on the audiotape will be changed to pseudonyms when the recordings are transcribed. Once transcribed and the research findings are complete, the recordings will be deleted. The transcripts will be kept for 3 years.

Disclosure of Alternative Procedures: This is not applicable to you for this research study.

Data Confidentiality or Anonymity: All data will be maintained as confidential (if collecting identifiable data, i.e., audio/video recordings) and no identifying information such as your name will appear in any publication or presentation of the data. The data will be confidential. I will be conducting in person interviews with you; however, I will use pseudonyms, and only I will know the match between pseudonym and your identity and the schools.

Storage of Data and Data Retention Period: Paper data will be stored in a locked filing cabinet in the researcher's office for three years and then will be shredded. The data will also be entered into a software program and stored on the researcher's password-protected computer for three years and then deleted. Only members of the research team will have access to the data.

Risks or Discomforts: There are no anticipated risks or discomforts for you with this study.

PERCEPTIONS OF TEACHER EVALUATION MODELS

Who to Contact Should You Experience Any Negative Effects from Participating in this Study:

Should you experience negative effects, please contact state counseling services provider for an appointment or for a recommendation for services with another provider.

Benefits: There are no perceived benefits for your participation in this study.

Voluntary Participation: Your participation in this study is completely voluntary and you are free to withdraw your permission at any time for any reason without penalty or prejudice from the investigator. Please feel free to ask any questions of the investigator before signing this form and at any time during the study.”

IRB Contact Information: For your rights as a participant, you may contact the following: For questions about your rights as a participant, please contact the Director, Office of Research Integrity, Ball State University, Muncie, IN 47306, (765) 285-5070 or at irb@bsu.edu.

Study Title: A Comparison of Indiana Administrator and Teacher Perceptions of Teacher Evaluation Models

Consent

I, _____, agree to participate in this research project entitled, A Comparison of Indiana Administrator and Teacher Perceptions of Teacher Evaluation Models. I have had the study explained to me and my questions have been answered to my satisfaction. I have read the description of this project and give my consent to participate. I understand that I will receive a copy of this informed consent form to keep for future reference.

To the best of my knowledge, I meet the inclusion/exclusion criteria for participation (described on the previous page) in this study.

Participant's Signature

Date

Researcher Contact Information

Principal Investigator:

Bret L. Daghe, Doctoral Student
Department of Educational Leadership
Ball State University
Muncie, IN 47306
Telephone: (765) 285-8490
Email: bldaghe@bsu.edu

Faculty Supervisor:

Dr. Kendra Lowery
Department of Educational Leadership
Ball State University
Muncie, IN 47306
Telephone: (765) 285-8490
Email: kplowery@bsu.edu